

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 21:41:11 ; Search time 9384 Seconds  
(without alignments)  
11230.094 Million cell updates/sec

Title: US-09-942-429A-6

Perfect score: 2576

Sequence: 1 gaattccgcagagatccagc.....aaaaaaaaaacggaattc 2576

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBankl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2576	100.0	2576	6	AX336419	AX336419 Sequence
2	2576	100.0	2576	6	AX409491	AX409491 Sequence
3	2576	100.0	2576	9	HUMCTTFAOH	L04751 Human cytoC
4	2417.2	93.8	2470	9	S67580	S67580 CYP4A11=fat
5	2394	92.9	2469	9	S67581	S67581 CYP4A11=fat
6	2342.6	90.9	2382	9	HUMFAOH	D26481 Homo sapien
7	1625.8	63.1	1763	6	AX336099	AX336099 Sequence
8	1625.8	63.1	1763	6	AX336427	AX336427 Sequence
9	1625.8	63.1	1763	6	AX336709	AX336709 Sequence
10	1625.8	63.1	1763	6	AX408963	AX408963 Sequence
11	1625.8	63.1	1763	9	HUMOMHY	D13705 Human mRNA
12	1389.4	53.9	1833	9	BC041158	BC041158 Homo sapi
13	1241.4	48.2	1790	4	RABCTP4A6	M28656 Rabbit cyto
14	1219	47.3	2191	4	RABCTP4A7	M29531 Rabbit cyto
15	1197.2	46.5	2377	4	OCPP450KD	X57209 Rabbit mRNA
16	1181.2	45.9	2356	4	RABCTP4A5	M28655 Rabbit cyto
17	1179.2	45.8	2929	4	RABCTP2	J02818 Rabbit cyto
18	1137.8	44.2	2400	4	SSC278474	AL278474 Sus scrofa
19	1132.8	44.0	2376	4	RABCTP4A7	M29530 Rabbit cyto
20	1127.4	43.8	1694	4	RABCTP4A7	M28657 Rabbit cyto
21	1122.2	43.6	2455	4	AF384031	AF384031 Sus scrofa
22	1101.8	42.8	123921	9	AL731892	AL731892 Human DNA
23	1092.6	42.4	14995	9	AF208532	AF208532 Homo sapi
24	1091.2	42.4	2376	10	AK098123	AK098123 Mus muscu
25	1091	42.4	122432	2	AL390073	AL390073 Homo sapi
26	1091	42.4	193358	9	AL135960	AL135960 Human DNA
27	1091	42.4	193471	9	HSAL31016	AJ131016 Homo sapi
28	1083.6	42.1	1945	10	BC025936	BC025936 Mus muscu
29	1083.6	42.1	2382	10	BC014721	BC014721 Mus muscu
30	1083.6	42.1	2431	10	BC026582	BC026582 Mus muscu
31	1076.8	41.8	1931	10	BC033924	BC033924 Mus muscu
32	1072.6	41.6	2357	10	BC031141	BC031141 Mus muscu
33	1059.2	41.1	1957	10	RATCTP450B	M37828 Rat cytochr
34	1050.6	40.8	1515	4	SSC318096	AJ318096 Sus scrofa
35	1050.6	40.8	1515	4	SSC318097	AJ318097 Sus scrofa
36	1022.4	39.7	2134	10	RNCTY452	X07259 Rat mRNA fo
37	1022.4	39.7	2462	6	AX401946	AX401946 Sequence
38	1022.4	39.7	2462	10	RATCTP4A	M14972 Rat cytochr
39	1012.8	39.3	2081	10	AK098101	AK098101 Mus muscu
40	1012.2	39.3	1887	10	MMCTP4A	X69296 M.musculus
41	1012	39.3	2077	10	AB018421	AB018421 Mus muscu
42	1012	39.3	2083	10	BC010747	BC010747 Mus muscu
43	1012	39.3	2085	10	BC051049	BC051049 Mus muscu
44	979.6	38.0	2140	10	RATIVAA	M33936 Rat CYP4a 1
45	968.8	37.6	2491	10	AK098088	AK098088 Mus muscu

ALIGNMENTS

RESULT 1	AX336419	2576 bp	DNA	linear	PAT 09-JAN-2002
LOCUS	AX336419	Sequence 6928 from Patent WO0194629.			
DEFINITION	AX336419				
ACCESSION	AX336419				
VERSION	AX336419.1	GI:18127138			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,				
AUTHORS	Horrikan,S., Soppet,D.R. and Weaver,Z.				
TITLE	Cancer gene determination and therapeutic screening using signature				

JOURNAL Patent: WO 0194629-A 6928 13-DEC-2001;  
Avalon Pharmaceuticals (US)  
FEATURES Location/Qualifiers  
source 1. 2576  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 572 a 793 c 583 g 628 t  
ORIGIN

Query Match 100.0%; Score 2576; DB 6; Length 2576;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATCCGAGAGATCCAGCAGGTGCTGACCATGAGTGTCTGTGCTGAGCCCCAGCA 60  
DB 1 GAATCCGAGAGATCCAGCAGGTGCTGACCATGAGTGTCTGTGCTGAGCCCCAGCA 60  
QY 61 GACTCCTGGGTGATGTCTCTGGAATCTTCCAAGCGGCTCCCTGCTCATTTCTGCTTGC 120  
DB 61 GACTCCTGGGTGATGTCTCTGGAATCTTCCAAGCGGCTCCCTGCTCATTTCTGCTTGC 120  
QY 121 TGTGATCAAGGAGTTCAGCTCTACTGCAAGGCAAGTGTGCTCAAAAGCCCTCCAGC 180  
DB 121 TGTGATCAAGGAGTTCAGCTCTACTGCAAGGCAAGTGTGCTCAAAAGCCCTCCAGC 180  
QY 181 AGTCCCGTGCCTCCCTCCCACTGGCTCTTGGGGCAATCCAGAGCTCCAAAGAGACC 240  
DB 181 AGTCCCGTGCCTCCCTCCCACTGGCTCTTGGGGCAATCCAGAGCTCCAAAGAGACC 240  
QY 241 AGAGCTACAAGGATTCAGAAATGGGTGAGACATTCCCAAGTCTGTCTCATTTGGC 300  
DB 241 AGAGCTACAAGGATTCAGAAATGGGTGAGACATTCCCAAGTCTGTCTCATTTGGC 300  
QY 301 TATGGGAGGCAAAAGTTCTGTCTCAGCTCTAATGACCTGACTATATGAAGTGATTCTGG 360  
DB 301 TATGGGAGGCAAAAGTTCTGTCTCAGCTCTAATGACCTGACTATATGAAGTGATTCTGG 360  
QY 361 GGAGATCAGACCCGAAATCCCATGTTCTCAAGATTCTTGCTCATGATTGGGTACG 420  
DB 361 GGAGATCAGACCCGAAATCCCATGTTCTCAAGATTCTTGCTCATGATTGGGTACG 420  
QY 421 GCTTGCTCCTGTGAATGGGAGACATGTTCCAGCATGACGAGTGTGACCCAGCCT 480  
DB 421 GCTTGCTCCTGTGAATGGGAGACATGTTCCAGCATGACGAGTGTGACCCAGCCT 480  
QY 481 TCCACTATGACATCTCTGAAGCCCTATGTGGGCTCATGGCAGACTCTGTACGATGC 540  
DB 481 TCCACTATGACATCTCTGAAGCCCTATGTGGGCTCATGGCAGACTCTGTGTATGC 540  
QY 541 TGGACAAATGGGAAGAGCTCCTTGGCCAGAGATTCCTTGAAGTCTTTCAAGCAGTCT 600  
DB 541 TGGACAAATGGGAAGAGCTCCTTGGCCAGAGATTCCTTGAAGTCTTTCAAGCAGTCT 600  
QY 601 CCTTGATGACCTTGACACCATCATGAGTGTGCTTCAAGCATCAGGAGCATCCAGG 660  
DB 601 CCTTGATGACCTTGACACCATCATGAGTGTGCTTCAAGCATCAGGAGCATCCAGG 660  
QY 661 TGGACAGGAATTCAGTCTCTACATACAGGCGCAATAGTACCTGAACAACCTGTTTTTT 720  
DB 661 TGGACAGGAATTCAGTCTCTACATACAGGCGCAATAGTACCTGAACAACCTGTTTTTT 720  
QY 721 CCCGTGTGAGGAATGCTTTTCAACAGAAATGACCATCTACAGCCTGCTGCGCC 780  
DB 721 CCCGTGTGAGGAATGCTTTTCAACAGAAATGACCATCTACAGCCTGCTGCGCC 780  
QY 781 GCTGACACACCGCGCTGCCAGCTGGCCCATCAGCACACAGACCAAGTATCCAACTGA 840  
DB 781 GCTGACACACCGCGCTGCCAGCTGGCCCATCAGCACACAGACCAAGTATCCAACTGA 840  
QY 841 GGAAGGCTCAACTACAGAAAGGGGGAGCTGGAAGATCAAGAGGAAGGCAATTTGG 900  
DB 841 GGAAGGCTCAACTACAGAAAGGGGGAGCTGGAAGATCAAGAGGAAGGCAATTTGG 900

DB 841 GGAAGGCTCAACTACAGAAAGGGGGAGCTGGAAGATCAAGAGGAAGGCAATTTGG 900  
QY 901 ATTTCTGATATTCCTCTCTTGGCCAAAATGAGAAATGGAGATCTTGTACAGACAGG 960  
DB 901 ATTTCTGATATTCCTCTCTTGGCCAAAATGAGAAATGGAGATCTTGTACAGACAGG 960  
QY 961 ACCTCCTGTGAGGTGACACGTTCAATGTTGAGGGCCACGACCAAGCAGTGGGA 1020  
DB 961 ACCTCCTGTGAGGTGACACGTTCAATGTTGAGGGCCACGACCAAGCAGTGGGA 1020  
QY 1021 TCTCTGATCTCTATAGTCTTGGCCACACCCCAAGCATCAGAGAGGTGCGGAGG 1080  
DB 1021 TCTCTGATCTCTATAGTCTTGGCCACACCCCAAGCATCAGAGAGGTGCGGAGG 1080  
QY 1081 AGATCCACAGCTCTCGGTGATGAGGCTCCATCACCTGGAACCACTGAGCAGATGC 1140  
DB 1081 AGATCCACAGCTCTCGGTGATGAGGCTCCATCACCTGGAACCACTGAGCAGATGC 1140  
QY 1141 CCTACACCAACATGTGCATTAAAGAGGCACTGAGGCTTACCCACCGGTGCCAGCATTG 1200  
DB 1141 CCTACACCAACATGTGCATTAAAGAGGCACTGAGGCTTACCCACCGGTGCCAGCATTG 1200  
QY 1201 GCAGAGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCTTGCCAAAGTATCA 1260  
DB 1201 GCAGAGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCTTGCCAAAGTATCA 1260  
QY 1261 TGTCTCTCTTCATTTATAGGCTTCAACCAACCCAAAGTGTGGCCAAACAGAGG 1320  
DB 1261 TGTCTCTCTTCATTTATAGGCTTCAACCAACCCAAAGTGTGGCCAAACAGAGG 1320  
QY 1321 TGTTCACCTTCCGTTTTCACCCGGTCTGCTCAACAAGGCTTCTTCTGCTG 1380  
DB 1321 TGTTCACCTTCCGTTTTCACCCGGTCTGCTCAACAAGGCTTCTTCTGCTGCTG 1380  
QY 1381 TCTCAGAGAGATCAAGGAATGCAATGGGAAACAATTTGCCATGACGAGCTGAAGTGG 1440  
DB 1381 TCTCAGAGAGATCAAGGAATGCAATGGGAAACAATTTGCCATGACGAGCTGAAGTGG 1440  
QY 1441 CCACGGCCCTGACCTGCTCCGCTTTCAGCTGCTGCTGATCCACAGGATCCCATCC 1500  
DB 1441 CCACGGCCCTGACCTGCTCCGCTTTCAGCTGCTGCTGATCCACAGGATCCCATCC 1500  
QY 1501 CCATTGACGACTGTGTGAATCCAAATGGAATCCACCTGCTCAGAGGCTCC 1560  
DB 1501 CCATTGACGACTGTGTGAATCCAAATGGAATCCACCTGCTCAGAGGCTCC 1560  
QY 1561 CTAACCTTGTGAGACAAGGACCAAGCTTGAAGGCTCCACCTGCGCTGCTTCTCCT 1620  
DB 1561 CTAACCTTGTGAGACAAGGACCAAGCTTGAAGGCTCCACCTGCGCTGCTTCTCCT 1620  
QY 1621 GACCCCGCTTCTGTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
DB 1621 GACCCCGCTTCTGTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
QY 1681 CCTTCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
DB 1681 CCTTCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
QY 1741 CAGGCTCCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
DB 1741 CAGGCTCCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
QY 1801 AAAGCTGAGTGTGGGAGAAAGCTGAGGCGGAGCTTGACATGCTGACATTAATGTAAGAG 1860  
DB 1801 AAAGCTGAGTGTGGGAGAAAGCTGAGGCGGAGCTTGACATGCTGACATTAATGTAAGAG 1860  
QY 1861 TCTTGAATCATGTCCAGATCCAGGCTTAACCCCTTGTGCTTGAACAACCAAGC 1920  
DB 1861 TCTTGAATCATGTCCAGATCCAGGCTTAACCCCTTGTGCTTGAACAACCAAGC 1920  
QY 1921 TCTGTGCTGAAGGTGGAAGGCTAACCTGACGACCATTAATCTAAGCCCGGGGCAATAAA 1980  
DB 1921 TCTGTGCTGAAGGTGGAAGGCTAACCTGACGACCATTAATCTAAGCCCGGGGCAATAAA 1980

QY	1981	CCCCCTGCTGGCTTGATAGATCCAGGGCTCGTGGCTCTGGAATGTGCTGCACTTGCTG	2040
Db	1981	CCCCCTGCTGGCTTGATAGATCCAGGGCTCGTGGCTCTGGAATGTGCTGCACTTGCTG	2040
QY	2041	GCTCCTTGCTCCTTGCTCTCCAGGATCAATTGTATCTTGAGTTAAAGAACCCTGCTC	2100
Db	2041	GCTCCTTGCTCCTTGCTCTCCAGGATCAATTGTATCTTGAGTTAAAGAACCCTGCTC	2100
QY	2101	CATTATCTCAAGTAACAGACAGATGCTAAACCGTCACAGCTGTAAATTGTGTCTTAAT	2160
Db	2101	CATTATCTCAAGTAACAGACAGATGCTAAACCGTCACAGCTGTAAATTGTGTCTTAAT	2160
QY	2161	GCAACATGCCCTTTGCAACCCACCCCATTCACCAACCTGTTCTTTGTTGATCAACA	2220
Db	2161	GCAACATGCCCTTTGCAACCCACCCCATTCACCAACCTGTTCTTTGTTGATCAACA	2220
QY	2221	ATAAATAATCTGCACTTCCAGAGCTCGGGCCTTCAGAGCCTCCATCCTTAGCTTGGCG	2280
Db	2221	ATAAATAATCTGCACTTCCAGAGCTCGGGCCTTCAGAGCCTCCATCCTTAGCTTGGCG	2280
QY	2281	CCCTGAGCCCACTTCTCTCTCAAACTGTCTTTCTCACTGCTTGACTCTGCCGACTT	2340
Db	2281	CCCTGAGCCCACTTCTCTCTCAAACTGTCTTTCTCACTGCTTGACTCTGCCGACTT	2340
QY	2341	TGTCACCCCCCAGACCTGTGTGGGTCTGAACACCCCAACATCCCTGAATCTCCACCA	2400
Db	2341	TGTCACCCCCCAGACCTGTGTGGGTCTGAACACCCCAACATCCCTGAATCTCCACCA	2400
QY	2401	CCTCCAAACTCCTGCTGCCCTCCAGACTGTCTGCCATACACCTGTCTCCTTCTTCT	2460
Db	2401	CCTCCAAACTCCTGCTGCCCTCCAGACTGTCTGCCATACACCTGTCTCCTTCTTCT	2460
QY	2461	GCCTGCTGTCTGTTCCTAATTAAGTTTCTATTACTGCTGTATAAATACTACAAATC	2520
Db	2461	GCCTGCTGTCTGTTCCTAATTAAGTTTCTATTACTGCTGTATAAATACTACAAATC	2520
QY	2521	TCAAGTATTTTAAATAAAAAATTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT	2576
Db	2521	TCAAGTATTTTAAATAAAAAATTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT	2576

RESULT 2	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	Source	BASE COUNT	ORIGIN	Query Match	Best Local Similarity	Matches 2576; Conservative	Score 2576; DB 6; Length 2576; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;	
AX409491	AX409491	Sequence 2138 from Patent WO0229103.	AX409491	AX409491.1	GI:21442196	Homo sapiens (human)	Homo sapiens	Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.	Gene expression profiles in liver cancer	Patent: WO 0229103-A 2138 11-APR-2002;	GENE LOGIC INC (US)	Location/Qualifiers	1. .2576	572 a	793 c	583 g	628 t	100.0%;	100.0%;	0;
												/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /note="EMBL/GenBank Accession No. L04751"								

Db	1	GAAATTCGGCAGAGATCCAGCAGAGTGCTGCACCCATGAGTGTCTCTGTGCTGAGCCCCCAGCA	60
QY	61	GACTCCTGGGTGATGTCCTGGAATCCTCCAAAGCGGCTCCCTGCTCATTTCTGCTTGC	120
Db	61	GACTCCTGGGTGATGTCCTGGAATCCTCCAAAGCGGCTCCCTGCTCATTTCTGCTTGC	120
QY	121	TGCTGATCAAGGCAAGTTCACTCTACCTGCACAGGCAATGGCTGCTCAAGCCCTCCAGC	180
Db	121	TGCTGATCAAGGCAAGTTCACTCTACCTGCACAGGCAATGGCTGCTCAAGCCCTCCAGC	180
QY	181	AGTTCCCGTGCCCTCCCTCCACTGCGCTTTGGGGCATCCAGAGCTCCCAAGAGACC	240
Db	181	AGTTCCCGTGCCCTCCCTCCACTGCGCTTTGGGGCATCCAGAGCTCCCAAGAGACC	240
QY	241	AGAGCTACAACGGATTCAGAAATGGGTGAGACATTTCCAAAGTGCCTGCTCATTTGC	300
Db	241	AGAGCTACAACGGATTCAGAAATGGGTGAGACATTTCCAAAGTGCCTGCTCATTTGC	300
QY	301	TATGGGAGGCAAGTTGCTGTCAGCTCTATGACCTGACTATATGAAGGTGATTCGG	360
Db	301	TATGGGAGGCAAGTTGCTGTCAGCTCTATGACCTGACTATATGAAGGTGATTCGG	360
QY	361	GGAGATCAGACCCGAAATCCCATGTTCTCTACAGATTCTGGCTCCATGGATTGGGTACG	420
Db	361	GGAGATCAGACCCGAAATCCCATGTTCTCTACAGATTCTGGCTCCATGGATTGGGTACG	420
QY	421	GCTTGCTCCTGTTGAATGGGAGACATGTTCCAGATCGACGGATGCTGACCCAGCT	480
Db	421	GCTTGCTCCTGTTGAATGGGAGACATGTTCCAGATCGACGGATGCTGACCCAGCT	480
QY	481	TCCACTATGACATCCTGAAGCCCTATGTTGGGGCTCATGGCAGACTCTGTACGAGTATGC	540
Db	481	TCCACTATGACATCCTGAAGCCCTATGTTGGGGCTCATGGCAGACTCTGTACGAGTATGC	540
QY	541	TGGACAAATGGGAGAGACTCTTGGCCAGGATTCCTCTGAGGCTTTTACGACAGTCT	600
Db	541	TGGACAAATGGGAGAGACTCTTGGCCAGGATTCCTCTGAGGCTTTTACGACAGTCT	600
QY	601	CCTTGATGACCCTGGAACACCATCATGAGTGTGCTTACGCCATCAGGGCAGCATCCAGG	660
Db	601	CCTTGATGACCCTGGAACACCATCATGAGTGTGCTTACGCCATCAGGGCAGCATCCAGG	660
QY	661	TGGACAGGAATTCAGTCTTACATACAGGCCAATGAGCTGAACAACCTGGTTTTTT	720
Db	661	TGGACAGGAATTCAGTCTTACATACAGGCCAATGAGCTGAACAACCTGGTTTTTT	720
QY	721	CCCGTGTAGGAATGCTTTTACCAAGATGACACCACTTAACGCTGACCTGTGCTGCGC	780
Db	721	CCCGTGTAGGAATGCTTTTACCAAGATGACACCACTTAACGCTGACCTGTGCTGCGC	780
QY	781	GCTGAGACACCGCGCTGCGCAGCTGGCCCATGACACACAGACCAAGTGATCCAACTGA	840
Db	781	GCTGAGACACCGCGCTGCGCAGCTGGCCCATGACACACAGACCAAGTGATCCAACTGA	840
QY	841	GGAAGCTCAACTACAGAAGAGGGGAGCTGAGAAATCAAGAGGAAGGCAATTTGG	900
Db	841	GGAAGCTCAACTACAGAAGAGGGGAGCTGAGAAATCAAGAGGAAGGCAATTTGG	900
QY	901	ATTTTCTGGATATCTCTCTTGGCCAAATGAGAAATGGGAGCATCTTGTCAAGCAAGG	960
Db	901	ATTTTCTGGATATCTCTCTTGGCCAAATGAGAAATGGGAGCATCTTGTCAAGCAAGG	960
QY	961	ACCTCCGTGCTGAGGTGACACGTTCAATGTTTGAAGGGCAGCAGACCAAGCCAGTGGGA	1020
Db	961	ACCTCCGTGCTGAGGTGACACGTTCAATGTTTGAAGGGCAGCAGACCAAGCCAGTGGGA	1020
QY	1021	TCTCCTGATCCTATGCTCTGGCCACACACCCCAAGCATCAGAGAGGTGCGGGAGG	1080
Db	1021	TCTCCTGATCCTATGCTCTGGCCACACACCCCAAGCATCAGAGAGGTGCGGGAGG	1080
QY	1081	AGATCCACAGCCTCTGGGTGATGAGCCTTCATCACTGGAAACCACTGGACCAAGATGC	1140
Db	1081	AGATCCACAGCCTCTGGGTGATGAGCCTTCATCACTGGAAACCACTGGACCAAGATGC	1140

OY 1141 CCTACACCACTGTCATTAAAGAGGCACTGAGGCTCTACCCACCGGTGCGGCAATTG 1200  
DB 1141 CCTACACCACTGTCATTAAAGAGGCACTGAGGCTCTACCCACCGGTGCGGCAATTG 1200  
OY 1201 GCAGAGAGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCTTGCCCAAGGTATCA 1260  
DB 1201 GCAGAGAGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCTTGCCCAAGGTATCA 1260  
OY 1261 TGGTCTCTCTCATTATTAAGGCTTCAACAACCAAAAGTGGCCCAACCCAGAGG 1320  
DB 1261 TGGTCTCTCTCATTATTAAGGCTTCAACAACCAAAAGTGGCCCAACCCAGAGG 1320  
OY 1321 TGTTCACCCCTTCCGTTTTCACCGGGTCTGCTCAACACAGCGCTTTCCTGCGCT 1380  
DB 1321 TGTTCACCCCTTCCGTTTTCACCGGGTCTGCTCAACACAGCGCTTTCCTGCGCT 1380  
OY 1381 TCTCAGAGAGATCAAGAACTGCAATTGGGAAACAATTTGCCATGAACGAGCTGAAGTGG 1440  
DB 1381 TCTCAGAGAGATCAAGAACTGCAATTGGGAAACAATTTGCCATGAACGAGCTGAAGTGG 1440  
OY 1441 CCACGGCCCTGACCTGCTCCGCTTTCAGCTGCTGATCCCAAGGATCCCATCC 1500  
DB 1441 CCACGGCCCTGACCTGCTCCGCTTTCAGCTGCTGATCCCAAGGATCCCATCC 1500  
OY 1501 CCATTGACAGACTGTGTGAATCCAAAATGGAATCCACCTGCTCAGAGGCTCC 1560  
DB 1501 CCATTGACAGACTGTGTGAATCCAAAATGGAATCCACCTGCTCAGAGGCTCC 1560  
OY 1561 CTAACTCTGTGAAGCAAGGACAGCTTGAAGGCTCCACCTGCGCTCTGCTTCT 1620  
DB 1561 CTAACTCTGTGAAGCAAGGACAGCTTGAAGGCTCCACCTGCGCTCTGCTTCT 1620  
OY 1621 GACCCCGCTTCTGTCCTCTCTGTCGCTATCTCTGTTTCTGTCGCCACCTTC 1680  
DB 1621 GACCCCGCTTCTGTCCTCTCTGTCGCTATCTCTGTTTCTGTCGCCACCTTC 1680  
OY 1681 CCTTCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
DB 1681 CCTTCTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740  
OY 1741 CCAGGCTCCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
DB 1741 CCAGGCTCCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
OY 1801 AAAGCTGAGTGTGGGAGAGCTGAGGCGGAGCTGATGTCTGACATATGTAAAGAG 1860  
DB 1801 AAAGCTGAGTGTGGGAGAGCTGAGGCGGAGCTGATGTCTGACATATGTAAAGAG 1860  
OY 1861 TCTTGAATCATGTCCAGATCCAGAGGCTTAACCCCTTGTGGCTTTGGAACAACCAAGC 1920  
DB 1861 TCTTGAATCATGTCCAGATCCAGAGGCTTAACCCCTTGTGGCTTTGGAACAACCAAGC 1920  
OY 1921 TCTGTGCTGAAGGGGTGAAGGCTACCTGACGCACTAATCTAAGCCCGGGGCAATAAA 1980  
DB 1921 TCTGTGCTGAAGGGGTGAAGGCTACCTGACGCACTAATCTAAGCCCGGGGCAATAAA 1980  
OY 1981 CCCCTGCTGCTTGAATGAATCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040  
DB 1981 CCCCTGCTGCTTGAATGAATCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040  
OY 2041 GCTCCTGCTCCTTGTCTCCAGGATCAATTTGATCTTGAAGTAAAGAACTGCTCTC 2100  
DB 2041 GCTCCTGCTCCTTGTCTCCAGGATCAATTTGATCTTGAAGTAAAGAACTGCTCTC 2100  
OY 2101 CATTATCTCAAGTAAACAGAGAGATGCTAAACCGTCAAGCTGTAATTTGTGCTTAAT 2160  
DB 2101 CATTATCTCAAGTAAACAGAGAGATGCTAAACCGTCAAGCTGTAATTTGTGCTTAAT 2160  
OY 2161 GCAACATGCTCTTTCGACCCCACTTCTCAACCACTGTTCTTGTGTTGATCAACA 2220  
DB 2161 GCAACATGCTCTTTCGACCCCACTTCTCAACCACTGTTCTTGTGTTGATCAACA 2220

OY 2221 ATAAATATCTGCACTTCCAGAGCTCGGGGCTTACAGGCTCCATCTTAGCTTGGCG 2280  
DB 2221 ATAAATATCTGCACTTCCAGAGCTCGGGGCTTACAGGCTCCATCTTAGCTTGGCG 2280  
OY 2281 CCCTGAGACCACTTCTCTCTCAAACTGCTTTTCTCACTGCTTGACTCTGCCGACTT 2340  
DB 2281 CCCTGAGACCACTTCTCTCTCAAACTGCTTTTCTCACTGCTTGACTCTGCCGACTT 2340  
OY 2341 TGTCAACCCCAAGACTGTGTTGGGTGTAACACCCCAACATCCCTGAATCTCCACCCA 2400  
DB 2341 TGTCAACCCCAAGACTGTGTTGGGTGTAACACCCCAACATCCCTGAATCTCCACCCA 2400  
OY 2401 CCTCCCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460  
DB 2401 CCTCCCAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460  
OY 2461 GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
DB 2461 GCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520  
OY 2521 TCACTGATTTTAAATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTC 2576  
DB 2521 TCACTGATTTTAAATTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATTC 2576  
  
RESULT 3  
HUMCYTFAOH 2576 bp mRNA linear PRI 31-DEC-1994  
LOCUS Human cytochrome p-450 4A (CYP4A) mRNA, complete cds.  
DEFINITION L04751  
VERSION L04751.1, GI:181396  
KEYWORDS cytochrome P450; cytochrome P450 4A; fatty acid omega-hydroxylase.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
  
REFERENCE  
AUTHORS Palmer, C.N., Richardson, T.H., Griffin, K.J., Hsu, M.H.,  
Muerhoff, A.S., Clark, J.E. and Johnson, E.F.  
TITLE Characterization of a cDNA encoding a human kidney, cytochrome  
P-450 4A fatty acid omega-hydroxylase and the cognate enzyme  
expressed in Escherichia coli  
JOURNAL Blochim. Biophys. Acta 1172 (1-2), 161-166 (1993)  
MEDLINE 93176801  
PUBMED 7679927  
COMMENT Original, source text: Homo sapiens adult kidney cDNA to mRNA.  
FEATURES  
source  
1..2576  
location/qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="kidney"  
/dev\_stage="adult"  
1..2576  
/gene="CYP4A"  
33..1592  
/gene="CYP4A"  
/function="fatty acid omega-hydroxylase"  
/standard\_name="CYP4A"  
/evidence="experimental"  
/product="cytochrome P450"  
/protein\_id="AA58436.1"  
/db\_xref="GI:181397"  
/translation="MSVSVLSPSRLGDVSGILQAASLLILLILIKAVOLYLHROWL  
LKAIOFPCEPSPHMLFGHIOLEQDOELQRIQWVETFPSPACPHMLWGKVRVQLYDP  
DYMKVILGRSDPRSHGSYRFLAPWIGYGLLLNGOTWFQHRRLMTPAFHYDILKPYVG  
LMADSVRVMLDKWBEILGODSPLRVFOHVSIMTLDTIMKCAFSGHSIQVDRNSQSYI  
QAISDLNLTVPSRVNAFHQNDTYSLSAGRWTHRACQLAHQHTDQVQLRKAQLOK  
EGLEIKIRKRLDVLIDILLILAKMENGSLSDKDLRAEVDTFMFEHDTTASGISWIL  
VALATPKHQRCEBEIHSLLGDGASITWNHLDMPTTMCIKALRLYPVPGIGRE  
LSTPVPDGRSLPKGIMVLSIYGLHNPKWPNBEVDFPRFAPGSAQSHAFLLPF  
SGSRNCIGKOPAMNELKVATALLRPELLDPFTRIPRIPIARLVLSKNGIHLRLR





Db 2041 GCTCCTTGCTCCTTGTCTCTCCAGATCAATTGTATCTTGAGTTAAAGAACCTGCTCTC 2100  
QY 2101 CATTAATCTCAAGTAACAGAGCAGATGCTAAACCGTCAACAGCTGTAAATGTGTCTTAAT 2160  
Db 2101 CATTAATCTCAAGTAACAGAGCAGATGCTAAACCGTCAACAGCTGTAAATGTGTCTTAAT 2160  
QY 2161 GCAACATGCCCCCTTTCGACCCACCCCATCTTCAACCACTGTTCTTTGTTGATCACC 2220  
Db 2161 GCAACATGCCCCCTTTCGACCCACCCCATCTTCAACCACTGTTCTTTGTTGATCACC 2220  
QY 2221 ATAATAATCTGCACTTCCAGAGCTCGGGGCTTCAAGCTCCATCTTAGCTTTGCGG 2280  
Db 2221 ATAATAATCTGCACTTCCAGAGCTCGGGGCTTCAAGCTCCATCTTAGCTTTGCGG 2280  
QY 2281 CCCTGAGACCACTTCTCTCTCAAACTGTCTTTCTCACTGCTTTGACTCTGCGGACTT 2340  
Db 2281 CCCTGAGACCACTTCTCTCTCAAACTGTCTTTCTCACTGCTTTGACTCTGCGGACTT 2340  
QY 2341 TGTCAACCCCAAGACCTGCTGTTGGGTCTGAACACCCCAACATCCCTGAATCTCCACCA 2400  
Db 2341 TGTCAACCCCAAGACCTGCTGTTGGGTCTGAACACCCCAACATCCCTGAATCTCCACCA 2400  
QY 2401 CCTCCCAAACTCTGCTGCTCCCTCCAGACTGTCTGCCATACACCTGTCTCTCTCTCT 2460  
Db 2401 CCTCCCAAACTCTGCTGCTCCCTCCAGACTGTCTGCCATACACCTGTCTCTCTCTCT 2460  
QY 2461 GCCTGCTGTCTGTTCTTATATTAGTTTCTTCTTCTGCTGTAATTAATATACCAATC 2520  
Db 2461 GCCTGCTGTCTGTTCTTATATTAGTTTCTTCTTCTGCTGTAATTAATATACCAATC 2520  
QY 2521 TCACTGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2576  
Db 2521 TCACTGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2576

RESULT 4

S67580 2470 bp mRNA linear PRI 26-FEB-1994  
LOCUS CYP4A11=fatty acid omega-hydroxylase [human, kidney, mRNA, 2470  
DEFINITION nt].

ACCESSION 867580  
VERSION 867580.1 GI:456995  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2470)  
AUTHORS Imaoka, S., Ogawa, H., Kimura, S. and Gonzalez, F.J.  
TITLE Complete cDNA sequence and cDNA-directed expression of CYP4A11, a  
fatty acid omega-hydroxylase expressed in human kidney  
JOURNAL DNA Cell Biol. 12 (10), 893-899 (1993)  
MEDLINE 94099889  
PUBMED 8274222

REMARK Genbank staff at the National Library of Medicine created this  
entry [NCBI gibbon 141537] from the original journal article.  
This sequence comes from Fig. 1.

FEATURES

source

1..2470  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..2470  
/gene="CYP4A11"  
/note="fatty acid omega-hydroxylase, CYP4A11"  
42..1601  
/gene="CYP4A11"  
/note="major fatty acid-metabolizing P450; This sequence  
comes from Fig. 1"  
/codon\_start=1  
/product="fatty acid omega-hydroxylase"  
/protein\_id="AAB29502.1"  
/db\_xref="GI:456995"

CDS

/translation="MSVSVLSPSRLLGDVSGILQAAALLILLILLIKAVQLYHROWL  
LKALQPPCPSPHMLFGHIOELQDQDELQRIQKWEVTPSPCPHMLGKVRVOLYDP  
DYMKVILGRSDPKSHGSRFLAPNGYGLLLNGQTFQHRMLTPAFHYDLKPYVG  
LMADSVRWLDMKEBLGDSPLFVQHSVLTMTDITMKCFASHQSIQVDRNSQSYI  
QALSDLANLVFSRVNAPFQNDITISLSAGRWTHRAQOLAHQHTDQVIOQRKQLOK  
EGLEKIKRKRDLFDILLAKENGSILSDKDLRAEVDFTMEGHDTTASGISWIL  
YALATPHKQERCBREIHSILGDGASITWNHLDQMPYTTMCKEALRLYPVPQIGRE  
LSTPVPFDPGRSLPKIMVLSIYGHNPKNVMPNPEVDFPSPFASQASHAFLPF  
SGSRNCIGKQFAMNELKVATALTLLRFELLPDPTRIPIPIARLVLSKNIGHLRLRR  
LPNPECDQL"

BASE COUNT 521 a 777 c 579 g 593 t  
ORIGIN

Query Match 93.8%; Score 2417.2; DB 9; Length 2470;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 2443; Conservative 0; Mismatches 3; Indels 4; Gaps 2;

QY 7 CGCAGAGATCCAGAGGTGCTGCACCATGAGTGTCTGTGCTGAGCCCGACAGACTCC 66  
Db 16 CTCAGAGATCCAGAGGTGCTGCACCATGAGTGTCTGTGCTGAGCCCGACAGACTCC 75  
QY 67 TGGGTGATGCTCTGGAATCTCCAGCGGCTCCCTGCTCATTCGCTTCTGCTGTA 126  
Db 76 TGGGTGATGCTCTGGAATCTCCAGCGGCTCCCTGCTCATTCGCTTCTGCTGTA 135  
QY 127 TCAAGGAGTTCACTCTACCTGCACAGGAGTGGCTGCTCAAGCCCTCCAGAGTTC 186  
Db 136 TCAAGGAGTTCACTCTACCTGCACAGGAGTGGCTGCTCAAGCCCTCCAGAGTTC 195  
QY 187 CGTGCCCTCCCTCCCACTGGCTCTTGGGGACATCCAGAGCTCCACAGACAGAGAGC 246  
Db 196 CGTGCCCTCCCTCCCACTGGCTCTTGGGGACATCCAGAGCTCCACAGACAGAGAGC 255  
QY 247 TACAACGAGTTCAGAAATGGGTGAGACATTCACAGTCCCTGCTCATTTGGCTATGG 306  
Db 256 TACAACGAGTTCAGAAATGGGTGAGACATTCACAGTCCCTGCTCATTTGGCTATGG 315  
QY 307 GAGCAAGTTCGTTCCAGCTCTATGACCTGACTATATGAAGGTGATTCGGGAGAT 366  
Db 316 GAGCAAGTTCGTTCCAGCTCTATGACCTGACTATATGAAGGTGATTCGGGAGAT 375  
QY 367 CAGACCCGAATCCATGCTTCTTCTACAGATTCTGCTCCATGATTTGGTAAGGCTTGC 426  
Db 376 CAGACCCGAATCCATGCTTCTTCTACAGATTCTGCTCCATGATTTGGTAAGGCTTGC 435  
QY 427 TCCTGTTGAATGGGACAGATGCTTCCAGCATTCACGATGCTGACCCCGAGCTTCCACT 486  
Db 436 TCCTGTTGAATGGGACAGATGCTTCCAGCATTCACGATGCTGACCCCGAGCTTCCACT 495  
QY 487 ATGACATCTGMAAGCCCTATGTGGGCTCATGGGAGACTCTGTACAGATGATGCTGACA 546  
Db 496 ATGACATCTGMAAGCCCTATGTGGGCTCATGGGAGACTCTGTACAGATGATGCTGACA 555  
QY 547 AATGGGAAGAGCTCTTGGCCAGAGATTCCCTGAGAGTCTTTACAGACGCTCTCTTGA 606  
Db 556 AATGGGAAGAGCTCTTGGCCAGAGATTCCCTGAGAGTCTTTACAGACGCTCTCTTGA 615  
QY 607 TGACCTTGACACCATCATGAAGTGTGCTTACGACCATCAGGGCAGCATCCAGGTGACA 666  
Db 616 TGACCTTGACACCATCATGAAGTGTGCTTACGACCATCAGGGCAGCATCCAGGTGACA 675  
QY 667 GGAATTTCTAGTCTTACATACAGGCCATTAGTGACCTGAACAACCTGTTTTCCTG 726  
Db 676 GGAATTTCTAGTCTTACATACAGGCCATTAGTGACCTGAACAACCTGTTTTCCTG 735  
QY 727 TGAGGAATGCTTTACAGAAATGACACCATCTACAGCTGACCTCTGCTGGCCGCTGGA 786  
Db 736 TGAGGAATGCTTTACAGAAATGACACCATCTACAGCTGACCTCTGCTGGCCGCTGGA 795  
QY 787 CACACCGCGCTGCGAGCTGGCCCATCAGACACAGCAAGTGTATCCAACTGAGGAAG 846  
Db 796 CACACCGCGCTGCGAGCTGGCCCATCAGACACAGCAAGTGTATCCAACTGAGGAAG 855

Qy 847 CTCACTACAGAGGAGGAGCTGAGAGATCAAGAGAGAGGCAATTGGATTTC 906  
Db 856 CTCAACTACAGAGGAGGAGAGCTGAGAGAGATCAAGAGAGAGGCAATTGGATTTC 915  
Qy 907 TGGATATCTCTCTTGGCCAAATGAGAGATGGAGCATTTGTCAACAAGACCTCC 966  
Db 916 TGGATATCTCTCTTGGCCAAATGAGAGATGGAGCATTTGTCAACAAGACCTCC 975  
Qy 967 GTGCTGAGGTGACACGTTCAATGTTGAGGGCCACGACACCAAGCAGTGGATCTCT 1026  
Db 976 GTGCTGAGGTGACACGTTCAATGTTGAGGGCCACGACACCAAGTGGATCTCT 1035  
Qy 1027 GGATCTCTATGCTCTGGCCACACACCCCAAGCATCAGAGAGGTGCCGGAGAGATCC 1086  
Db 1036 GGATCTCTATGCTCTGGCCACACACCCCAAGCATCAGAGAGGTGCCGGAGAGATCC 1095  
Qy 1087 ACAGCTCTGGGTGATGAGGCTTCATCACTGGAACCACTGGAACAGATGCCCTACA 1146  
Db 1096 ACAGCTCTGGGTGATGAGGCTTCATCACTGGAACCACTGGAACAGATGCCCTACA 1155  
Qy 1147 CCACCATGTCATTAAAGAGGCACTGAGGCTTACCCACCGGTGCCAGGCAATGGCAGAG 1206  
Db 1156 CCACCATGTCATTAAAGAGGCACTGAGGCTTACCCACCGGTGCCAGGCAATGGCAGAG 1215  
Qy 1207 AGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCCTGCCCCAAAGGTATCATGCTCC 1266  
Db 1216 AGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCCTGCCCCAAAGGTATCATGCTCC 1275  
Qy 1267 TCCTCTCCATTATGCGCTTCAACCAACCCAAAGTGTGGCCCAACCAAGAGTGTTC 1326  
Db 1276 TCCTCTCCATTATGCGCTTCAACCAACCCAAAGTGTGGCCCAACCAAGAGTGTTC 1335  
Qy 1327 ACCCTTTCGTTTGTGACCGGGTTTGTCTCAACACAGCCAGCTTTCCTGCCCTTCTCAG 1386  
Db 1336 ACCCTTTCGTTTGTGACCGGGTTTGTCTCAACACAGCCAGCTTTCCTGCCCTTCTCAG 1395  
Qy 1387 GAGGATCAAGAGAACTGCAATTGGGAAACAATTGCAATGAACGAGCTGAAGTGGCCACGG 1446  
Db 1396 GAGGATCAAGAGAACTGCAATCGGAAACAATTGCAATGAACGAGCTGAAGTGGCCACGG 1455  
Qy 1447 CCCTGACCTGCTCCGCTTGTGAGCTGCTGCTGATCCACAGAGATCCCATCCCATTCG 1506  
Db 1456 CCCTGACCTGCTCCGCTTGTGAGCTGCTGCTGATCCACAGAGATCCCATCCCATTCG 1515  
Qy 1507 CACGACTTGTGTGAATCCAAAATGGAATCCACCTGCGTTCAGAGAGCTCCCTAAC 1566  
Db 1516 CACGACTTGTGTGAATCCAAAATGGAATCCACCTGCGTTCAGAGAGCTCCCTAAC 1575  
Qy 1567 CTTGTGAAGACAAGGACCAAGCTTTGAGGGCTCCACCTGCGTCTGTCTTCTGACCCC 1626  
Db 1576 CTTGTGAAGACAAGGACCAAGCTTTGAGGGCTCCACCTGCGTCTGTCTTCTGACCCC 1635  
Qy 1627 CGCTTCTGTCCTTCTCTGTCGCCCAATATCTGTTTCTGTCTGACCACTTCCCTTCT 1686  
Db 1636 CGCTTCTGTCCTTCTCTGTCGCCCAATATCTGTTTCTGTCTGACCACTTCCCTTCT 1695  
Qy 1687 TCCCACTGCTGCTGTCGCCCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1745  
Db 1696 TCCCACTGCTGCTGTCGCCCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1755  
Qy 1746 CTCCCTACCTGCTGTTCTACCTGTCTCTACCCACCTGTATCTTGTGGAGAAAAGC 1805  
Db 1756 CTCCCTACCTGCTGTTCTACCTGTCTCTACCCACCTGTATCTTGTGGAGAAAAGC 1815  
Qy 1806 TGAGTGTGGAGAGAGCTGAGGCGGAGCTTGCATGTCTGACATATGTAAAGAGTCTTG 1865  
Db 1816 TGAGTGTGGAGAGAGCTGAGGCGGAGCTTGCATGTCTGACATATGTAAAGAGTCTTG 1875  
Qy 1866 AATCATGTCCAGATCCAGGCTTAACCCCTTGTGACCTTGGAAACCAAGCTCTGT 1925  
Db 1876 AATCATGTCCAGATCCAGGCTTAACCCCTTGTGACCTTGGAAACCAAGCTCTGT 1935  
Qy 1926 GCTGAAGGTGGAAGGCTACCTGACGACCATATATTAAGCCGGGCATTAACCCCT 1985

Db 1936 GCTGAAGGTGGAAGGCTACCTGACGCAACATATTAAGCCGGGCATTAACCCCT 1995  
Qy 1986 CGTGCTGTGATAGATCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2045  
Db 1996 CGTGCTGTGATAGATCCAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2055  
Qy 2046 TTGCTCTGCTCTCCAGATCAATTGTATCTTGAATTAAGAACTGCTCTCATTA 2105  
Db 2056 TTGCTCTGCTCTCCAGATCAATTGTATCTTGAATTAAGAACTGCTCTCATTA 2115  
Qy 2106 TCTCAAGTACAGAGCAGATGCTAAACGCTCACAGCTGTAAATTGTGTCTTAATGCAAC 2165  
Db 2116 TCTCAAGTACAGAGCAGATGCTAAACGCTCACAGCTGTAAATTGTGTCTTAATGCAAC 2175  
Qy 2166 ATGCCCTTGCACCCACCCCATCTTCAACACCTGTTCTTGTGATCACCAATAAA 2225  
Db 2176 ATGCCCTTGCACCCACCCCATCTTCAACACCTGTTCTTGTGATCACCAATAAA 2232  
Qy 2226 TAATCTGACTTCCAGAGCTGGGGCTTCAACAGCTCCATCTTAGCTTGGCGCCCTG 2285  
Db 2233 TAATCTGACTTCCAGAGCTGGGGCTTCAACAGCTCCATCTTAGCTTGGCGCCCTG 2292  
Qy 2286 GACCACTTCTCTCTCAAACTGTCTTCTCACTGCTTGAATCTGCGGACTTGTCA 2345  
Db 2293 GACCACTTCTCTCTCAAACTGTCTTCTCACTGCTTGAATCTGCGGACTTGTCA 2352  
Qy 2346 CCCCCACGACTGTGTGGGTGTAACAACCCCAACATCCCTGAATCTCCACCACTCC 2405  
Db 2353 CCCCCACGACTGTGTGGGTGTAACAACCCCAACATCCCTGAATCTCCACCACTCC 2412  
Qy 2406 CAAACTCTGCTGCCCTCCAGACTGTCTGCCATACACCTGTCTCTTC 2455  
Db 2413 CAAACTCTGCTGCCCTCCAGACTGTCTGCCATACACCTGTCTCTTC 2462  
  
RESULT 5  
S67581 2469 bp mRNA linear PRI 26-FEB-1994  
LOCUS S67581  
DEFINITION CYP4A11=fatty acid omega-hydroxylase [human, kidney, mRNA Mutant,  
2469 nt]  
ACCESSION S67581  
VERSION S67581.1 GI:456997  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 2469)  
AUTHORS Imaoka, S., Ogawa, H., Kimura, S. and Gonzalez, F. J.  
TITLE complete cDNA sequence and cDNA-directed expression of CYP4A11, a  
fatty acid omega-hydroxylase expressed in human kidney  
JOURNAL DNA Cell Biol. 12 (10), 893-899 (1993)  
MEDLINE 94099889  
PUBMED 8274222  
REMARK GenBank staff at the National Library of Medicine created this  
entry [NCBI glibseq 141542] from the original journal article.  
This sequence comes from Fig. 1.  
FEATURES  
source  
1..2469  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..2469  
gene  
/gene="CYP4A11"  
/note="fatty acid omega-hydroxylase, CYP4A11v"  
42..1817  
CDS  
/gene="CYP4A11"  
/note="This sequence comes from Fig. 1; CYP4A11v"  
/codon\_start=1  
/product="fatty acid omega-hydroxylase"  
/protein\_id="AAB29503.1"  
/db\_xref="GI:456998"



/translation="MSVSVLSPSRLLGDVSGILQAAALLILLILLIKAVOLYLHRLWL  
LKALQOPPCPSHMLFGHIOELOQDLORIQKMBETFPSPACHMLMGKRVOLDP  
DYMKVILGRSDPKSHGSYRFLAPWIGYLLLLNGQTFQHRMLTPAFHYDILKPYVG  
LMADSVRWLDMKEBELGDSPLFVFOHVSIMTLDTIMKCAFSGSIQVDRNOSYI  
QAIISDLNLVFSRVNAPHONDITYSLTSAGRWTHRACQOLAHOHTDVIQLRKAQLOK  
EGLEKIKRKRHLDFLDILLAKMENGSLSDKDLRAVDTFMFEGHDTTASGISMTL  
YALATHPKHQRCEBBIHSLGDASITWNHLDOMPYYTMCIKELRLYPVPVGIERS  
LSTPVPFQGRSLPKGIMVLSIYGLHNPKVMPNBEVFPDRFAPGSAQSHAFILP  
SGSRNICGQFQFAMNBLKVATALTALFPELLPDPTRIPILARLVLSKIMESTCVSG  
SLTVKTRTSFEGHLPLPSCLPDRFCLPVCPPYPVCLPTPPSHLPAVPQACPBLSL  
HLSPGLPTCLSTCLLPTCISCWEXS"

BASE COUNT 521 a 780 c 577 g 591 t  
ORIGIN

Query Match 92.9%; Score 2394; DB 9; Length 2469;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 2435; Conservative 0; Mismatches 10; Indels 5; Gaps 3;

QY 7 CGCAGAGATCCAGCAGTGCTGCACACCATGAGTGTCTGTGCTGAGGCCCGACAGACTCC 66  
DB 16 CTCAGAGATCCAGCAGTGCTGCACACCATGAGTGTCTGTGCTGAGGCCCGACAGACTCC 75  
QY 67 TGGGTGATGTCTGGAATCTCCAGCGGCTCTGCTCATTTGCTTCTGCTGCTGA 126  
DB 76 TGGGTGATGTCTGGAATCTCCAGCGGCTCTGCTCATTTGCTTCTGCTGCTGA 135  
QY 127 TCAAGGAGTTCAGCTCTACCTGCACAGGAGTGCTGCTCAAGCCCTCCAGAGTTCC 186  
DB 136 TCAAGGAGTTCAGCTCTACCTGCACAGGAGTGCTGCTCAAGCCCTCCAGAGTTCC 195  
QY 187 CGTGCCCTCCCTCCCACTGGCTCTTGGGGACATCCAGAGCTCCAGAGCCAGAGAGC 246  
DB 196 CGTGCCCTCCCTCCCACTGGCTCTTGGGGACATCCAGAGCTCCAGAGCCAGAGAGC 255  
QY 247 TACAACGAGTTCAAAAATGGGTGAGACATTCACAGTGCTGTCTCATTTGGCTATGGG 306  
DB 256 TACAACGAGTTCAAAAATGGGTGAGACATTCACAGTGCTGTCTCATTTGGCTATGGG 315  
QY 307 GAGGCAAGTTCGTCTCAGCTCTATGACCTGACTATATGAAAGTGTCTGGGGAGAT 366  
DB 316 GAGGCAAGTTCGTCTCAGCTCTATGACCTGACTATATGAAAGTGTCTGGGGAGAT 375  
QY 367 CAGACCCGAATCCCATGGTCTCTACAGATTCCTGGCTCCATGATGGGTACGGCTTGC 426  
DB 376 CAGACCCGAATCCCATGGTCTCTACAGATTCCTGGCTCCATGATGGGTACGGCTTGC 435  
QY 427 TCCTGTGTAATGGGCAACATGGTTCAGCATCGACGAGATGCTGACCCCGACTTCCACT 486  
DB 436 TCCTGTGTAATGGGCAACATGGTTCAGCATCGACGAGATGCTGACCCCGACTTCCACT 495  
QY 487 ATGACATCTGAAAGCCCTATGTGGGCTCATGGGAGACTCTGTACGAGTGTGCTGAGCA 546  
DB 496 ATGACATCTGAAAGCCCTATGTGGGCTCATGGGAGACTCTGTACGAGTGTGCTGAGCA 555  
QY 547 AATGGGAAGAGCTCTTGCCAGAGATTCCCTCTGAGAGTCTTTTCAGCAGCTCTCTTGA 606  
DB 556 AATGGGAAGAGCTCTTGCCAGAGATTCCCTCTGAGAGTCTTTTCAGCAGCTCTCTTGA 615  
QY 607 TGAACCTGAGACACCATCATGAAGTGTGCTTCAAGCATCAAGGAGAGATCCAGGTGAGCA 666  
DB 616 TGAACCTGAGACACCATCATGAAGTGTGCTTCAAGCATCAAGGAGAGATCCAGGTGAGCA 675  
QY 667 GGAATTTCTAGTCTCTACATACAGGCCATTAGTGAAGCTGAAACAACCTGTTTTTTCCCGTG 726  
DB 676 GGAATTTCTAGTCTCTACATACAGGCCATTAGTGAAGCTGAAACAACCTGTTTTTTCCCGTG 735  
QY 727 TGAAGAATGCTTTTCAACCAAGATGACACCATCTACAGCTGACCTGTGCTGGCCGCTGGA 786  
DB 736 TGAAGAATGCTTTTCAACCAAGATGACACCATCTACAGCTGACCTGTGCTGGCCGCTGGA 795  
QY 787 CACACCGCGCTGCGCAGCTGCGCCATCAGCACACAGCAAGTGTCCAACTGAGGAAGG 846  
DB 796 CACACCGCGCTGCGCAGCTGCGCCATCAGCACACAGCAAGTGTCCAACTGAGGAAGG 855

QY 847 CTCACCTACAGAAAGAGGGGAGCTGAGAAAGATCAAGAGAGAGGCAATTGGATTTC 906  
DB 856 CTCACCTACAGAAAGAGGGGAGCTGAGAAAGATCAAGAGAGAGGCAATTGGATTTC 915  
QY 907 TGAATATCTCTCTTGGCCAAATGAGAAATGGAGCATCTTGTCAACAAGACCTCC 966  
DB 916 TGAATATCTCTCTTGGCCAAATGAGAAATGGAGCATCTTGTCAACAAGACCTCC 975  
QY 967 GTGCTGAGGTGAGACAGTTTATGTTGAGGGCCACAGACACAGCCAGTGGATCTCCT 1026  
DB 976 GTGCTGAGGTGAGACAGTTTATGTTGAGGGCCACAGACACAGCCAGTGGATCTCCT 1035  
QY 1027 GGAATCTCTATGCTCTGGCCACACACCCCAAGATCAAGAGAGTGCCTGGGAGAGATCC 1086  
DB 1036 GGAATCTCTATGCTCTGGCCACACACCCCAAGATCAAGAGAGTGCCTGGGAGAGATCC 1095  
QY 1087 ACAGCCTCTGGGTGATGAGAGCTTCATCACTGGAAACCACTGGACCAAGATGCCCTACA 1146  
DB 1096 ACAGCCTCTGGGTGATGAGAGCTTCATCACTGGAAACCACTGGACCAAGATGCCCTACA 1155  
QY 1147 CAACCATGTGATTAAGAGGACATGAGGCTTACCAACCGGTGACAGGCAATTGGCAGAG 1206  
DB 1156 CAACCATGTGATTAAGAGGACATGAGGCTTACCAACCGGTGACAGGCAATTGGCAGAG 1215  
QY 1207 AGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCTTGCCCAAGATATCATGTCC 1266  
DB 1216 AGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCTTGCCCAAGATATCATGTCC 1275  
QY 1267 TCCTCTCATTTATAGGCTTCAACCAACCCCAAAAGTGTGGCCCAACCCAGAGTGTTC 1326  
DB 1276 TCCTCTCATTTATAGGCTTCAACCAACCCCAAAAGTGTGGCCCAACCCAGAGTGTTC 1335  
QY 1327 ACCCTTCCGTTTGAACCGGGTCTGCTCAACAGACCAAGCTTCTGCTCTCTCAG 1386  
DB 1336 ACCCTTCCGTTTGAACCGGGTCTGCTCAACAGACCAAGCTTCTGCTCTCTCAG 1395  
QY 1387 GAGGATCAAGAACTGCAATTGGGAAACAATTGGCATGAACGAGCTGAAGGTGGCCACGG 1446  
DB 1396 GAGGATCAAGAACTGCAATTGGGAAACAATTGGCATGAACGAGCTGAAGGTGGCCACGG 1455  
QY 1447 CCCTGACCTGCTCCGCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1506  
DB 1456 CCCTGACCTGCTCCGCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1515  
QY 1507 CACGACTGTGTGAATTCACAAAATGGAATCCACTGCTCTCAGAGGCTCCCTTAACC 1566  
DB 1516 CACGACTGTGTGAATTC - AAAATGGAATCCACTGCTCTCAGAGGCTCCCTTAACC 1574  
QY 1567 CTGTGGAAGACAAGACAGACTTGAAGGCTCCCACTGCGCTGCTGCTGCTGCTGCTGCTGCT 1626  
DB 1575 CTGTGGAAGACAAGACAGACTTGAAGGCTCCCACTGCGCTGCTGCTGCTGCTGCTGCTGCT 1634  
QY 1627 CGCTTCTGCTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1686  
DB 1635 CGCTTCTGCTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1694  
QY 1687 TCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1745  
DB 1695 TCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1754  
QY 1746 CTCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1805  
DB 1755 CTCCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1814  
QY 1806 TGAAGTGTGGAGAGAGCTGAGGCGAGGCTTGCATGTCTGACATATGTAAAGAGTCTTG 1865  
DB 1815 TGAAGTGTGGAGAGAGCTGAGGCGAGGCTTGCATGTCTGACATATGTAAAGAGTCTTG 1874  
QY 1866 AATCATGTCCAGATCCAGGGTCTAAACCCCTTGTGGCTTTGGAACACCAAGCTCTGT 1925  
DB 1875 AATCATGTCCAGATCCAGGGTCTAAACCCCTTGTGGCTTTGGAACACCAAGCTCTGT 1934

QY	1926	GCTGAAGGCTGAAAGGCTACCTCTGACGCAACCATATCTAAGCCGGGGCATAAAAACCCCT	1985
Db	1935	GCTGAAGGGGTGAAAGGCTACCTCTGACGCAACCATATCTAAGCCGGGGCATAAAAACCCCT	1994
QY	1986	CGTGGCTTGATAGAAATCCAGGGCTCGTGGCTCTGGAATGTCCTGACTTGTGCTGCC	2045
Db	1995	CGTGGCTTGATAGAAATCCAGGGCTCGTGGCTCTGGAATGTCCTGACTTGTGCTGCC	2054
QY	2046	TTGCTCCTTGTCTCTCCCAAGATCAATTGTATCTTGAAGTAAAGAACTGCTCTCCATTA	2105
Db	2055	TCGCTCCTTGTCTCTCCCAAGATCAATTGTATCTTGAAGTAAAGAACTGCTCTCCATTA	2114
QY	2106	TCTCAAGTAAACAGACAGATGCTAAACCGTCAACAGCTGTAAATGTGTGCTTAATGCAAC	2165
Db	2115	TCTCAAGTAAACAGACAGATGCTAAACCGTCAACAGCTGTAAATCATGTGCTTAATGCAAC	2174
QY	2166	ATGCCCTTTGACCCCAACCCCCCATTTCTCAACCACTGTCTTTGTTTGATCACCATAAA	2225
Db	2175	ATGCCCTTTGAC--CCCAATTTCTCAACCACTGTCTTTGTTTGATCACCATAAA	2231
QY	2226	TAATCTGCACTTCCAGAGCTCGGGCCCTTCAACAGCCCTCCATCCTTAAGCTTTGGCGCCCTG	2285
Db	2232	TAATCTGCACTTCCAGAGCTCGGGCCCTTCAACAGCCCTCCATCCTTAAGCTTTGGCGCCCTG	2291
QY	2286	GACCCCACTTCTCTCTCAAACTGTCTTTTCTCACTGCTTGACTCTGCCGGACTTTGTCA	2345
Db	2292	GACCCCACTTCTCTCTCAAACTGTCTTTTCTCACTGCTTGACTCTGCCGGACTTTGTCA	2351
QY	2346	CCCCCAAGACCTGTGTTGGGCTGAACAACCCCAACATCCCTGAATCTCCACCCACCTCC	2405
Db	2352	CCCCCAAGACCTGTGTTGGGCTGAACAACCCCAACATCCCTGAATCTCCACCCACCTCC	2411
QY	2406	CAAACTCCTGCTGCCCTCCAGACTGTCTGCCCATACACTGTCTCCTTC	2455
Db	2412	CAAACTCCTGCTGCCCTCCAGACTGTCTGCCCATACACTGTCTCCTTC	2461

```

RESULT 6
HUMPAOH
LOCUS      2382 bp      mRNA      linear      PRI 11-FEB-2003
DEFINITION Homo sapiens mRNA for fatty acids omega-hydroxylase, complete cds.
ACCESSION  D26481
VERSION    D26481.1  GI:536878
KEYWORDS
SOURCE
ORGANISM   Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 2382)
            Kawashima, H., Kusumose, E., Kikuta, Y., Kinoshita, H., Tanaka, S.,
            Yamamoto, S., Kishimoto, T. and Kusumose, M.
            Purification and cDNA cloning of human liver CYP4A fatty acid
            omega-hydroxylase
            J. Biochem. 116 (1), 74-80 (1994)
JOURNAL    MEDLINE
MEDLINE    95096029
PUBMED     7798189
REFERENCE  2 (bases 1 to 2382)
            Kikuta, Y.
            Direct Submission
            Submitted (11-JAN-1994) Yasushi Kikuta, Fukuyama University,
            Faculty of Engineering, Gakuenchou 1, Fukuyama, Hiroshima 729-0292,
            Japan (Tel:81-849-36-2111, Fax:81-849-36-2023)
JOURNAL
FEATURES
Source     1. 2382
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="HL24"
            /sex="male"
            /tissue_type="liver"
            /clone_fib="CLONTECH HL115b 25596"
            /dev_stage="37 years old"
            16. 1575

```

		/note="cytochrome P450H1L omega"
		/codon_start=1
		/product="fatty acids omega-hydroxylase"
		/protein_id="BAA05491.1"
		/db_xref="GI:994758"
		/translation="MSVSVLSPSRLGDVSGILQASLLILLILLIKAVQLYLHRWL LKALQFPCCPSHWLFGLIGELQDOQLRIQKWEVETPSPACPHLMWGKVRVOLYDP DYMKVILGRSDPKSHGSYRFLAPWYGILLNGOTWFQHRMLTPAFHYDLKPVG LMADSVRMLDKMBELLGDSPLEVFQHVSLMTLDTIMKCAFSSHOSIQVDNRSQYI OAI SDLNLI VFSRVRNAFHNDPTYSLSAGRWTHACOLAHQHDQVIQLRAQLOK EGELEKI KRKRHLDFLDILLAKMENGSI LSKDLRAEVDITMEGHDDTASGISWIL VALATHEKHGERCBEBIHSLDGASITWNHLDMQPYTMCIKEALRLYPVPVIGRE LSTPTPEPDGRSLPKGIMVLISLYGLHNPKWPNEVEFDPFRPAQSASHAFLPF SGSRNCIGKQFAMNELKVATLTLLRFELLDPTRIPPIARLVLSKNGIHLRLRR LPNPCEDKQL"
BASE COUNT	505 A 744 C 560 G 573 T	
ORIGIN		
Query Match	90.9%; Score 2342.6; DB 9; Length 2382;	
Best Local Similarity	99.5%; Pred. No. 0;	
Matches 2372; Conservative	0; Mismatches 9; Indels 4; Gaps 2;	
QY	18 AGCAGGTGCTGCACCATGAGTGTCCTGTGCTGAAGCCCCCAGCAGACTCTGGGTGATGTC	77
DB	1 AGCAGGTGCTGCACCATGAGTGTCCTGTGCTGAAGCCCCCAGCAGACTCTGGGTGATGTC	60
QY	78 TCTGAATCCTCCAAGCGGCTCCCTGCTCATTTCTGCTTCTGCTGTCATCAAGGCAGTT	137
DB	61 TCTGAATCCTCCAAGCGGCTCCCTGCTCATTTCTGCTTCTGCTGTCATCAAGGCAGTT	120
QY	138 CAGCTTACCTGCA CAGGAGTGGCTGCTCAAAGCCCTCCAGCAGTTCCCGTGCCCTCCC	197
DB	121 CAGCTTACCTGCA CAGGAGTGGCTGCTCAAAGCCCTCCAGCAGTTCCCGTGCCCTCCC	180
QY	198 TCCCACTGGCTCTTGGGACATCCAGAGCTTCCAACAAGACCAGAGCTTAACAAGGATT	257
DB	181 TCCCACTGGCTCTTGGGACATCCAGAGCTTCCAACAAGACCAGAGCTTAACAAGGATT	240
QY	258 CAGAATGGGTGAGACATTTCCCAAGTGCCTGCTCATTTGGCTATGGGAGGCCAAAGTT	317
DB	241 CAGAATGGGTGAGACATTTCCCAAGTGCCTGCTCATTTGGCTATGGGAGGCCAAAGTT	300
QY	318 CGTGTCCAGCTTATGACCCCTGACTATATGAAGTGATTCTGGGAGATGACACCCGAAA	377
DB	301 CGTGTCCAGCTTATGACCCCTGACTATATGAAGTGATTCTGGGAGATGACACCCGAAA	360
QY	378 TCCCATGGTTCCTACAGATTCCCTGGCTCCATGGATTGGGTACGGCTTGCTCTGTGAAT	437
DB	361 TCCCATGGTTCCTACAGATTCCCTGGCTCCATGGATTGGGTACGGCTTGCTCTGTGAAT	420
QY	438 GGGCAGACATGGTTCAGCATCGACGATGCTGACCCAGCCTTCCACTATGACATCCCTG	497
DB	421 GGGCAGACATGGTTCAGCATCGACGATGCTGACCCAGCCTTCCACTATGACATCCCTG	480
QY	498 AAGCCCTATGTGGGGCTCATGGCAGACTCTGTACGAGTGATGCTGGACAAATGGGAAGAG	557
DB	481 AAGCCCTATGTGGGGCTCATGGCAGACTCTGTACGAGTGATGCTGGACAAATGGGAAGAG	540
QY	558 CTCCTTGGCCAGGATTCCCTCTGAGAGTCTTTCAGCACGCTCTCTTATGACCCCTGGAC	617
DB	541 CTCCTTGGCCAGGATTCCCTCTGAGAGTCTTTCAGCACGCTCTCTTATGACCCCTGGAC	600
QY	618 ACCATCATGAAGTGTGCTTACGCCATCAAGGACGATCCAGGTGACAGGAATTCTCAG	677
DB	601 ACCATCATGAAGTGTGCTTACGCCATCAAGGACGATCCAGGTGACAGGAATTCTCAG	660
QY	678 TCCTACATACAGGCATTAGTGACCTGAACAACCTGGTTTTTCCCGTGTGAGGAATGCC	737
DB	661 TCCTACATACAGGCATTAGTGACCTGAACAACCTGGTTTTTCCCGTGTGAGGAATGCC	720
QY	738 TTTCACGAGAAATGACACCATCTACAGCCTGACCTCTGCTGGCCGCTGACACACCGCGCC	797
DB	721 TTTCACGAGAAATGACACCATCTACAGCCTGACCTCTGCTGGCCGCTGACACACCGCGCC	780

```
OY 798 TGCCAGTGGCCCATCAGACACAGACCAAGTATCCAACTGAGGAAGGCTCAACTACAG 857
|
|
|
DB 781 TGCCAGTGGCCCATCAGACACAGACCAAGTATCCAACTGAGGAAGGCTCAACTACAG 840
|
|
|
OY 858 AAGAGGGGGAGCTGAGAGATCAAGAGGAAGGCAATTGGATTTCTGATATCTC 917
|
|
|
DB 841 AAGAGGGGGAGCTGAGAGATCAAGAGGAAGGCAATTGGATTTCTGATATCTC 900
|
|
|
OY 918 CTCTTGGCCAAAATGAGAGATGGAGCATTTGTCAACAAGGACCTCCGTGAGGTG 977
|
|
|
DB 901 CTCTTGGCCAAAATGAGAGATGGAGCATTTGTCAACAAGGACCTCCGTGAGGTG 960
|
|
|
OY 978 GACACGTTCAATGTTGAGGGGCCAGACCAAGCAGTGGATCTCTGATCTCTAT 1037
|
|
|
DB 961 GACACGTTCAATGTTGAGGGGCCAGACCAAGCAGTGGATCTCTGATCTCTAT 1020
|
|
|
OY 1038 GCTCTGGCCACACACCCCAAGCATCAGAGAGGTGCCGGAGAGATCCACAGCCTCTG 1097
|
|
|
DB 1021 GCTCTGGCCACACACCCCAAGCATCAGAGAGGTGCCGGAGAGATCCACAGCCTCTG 1080
|
|
|
OY 1098 GGTGATGAGGCTTCATCACTTGAAACCACTGGAACAGATGCCCTACACCACTATGTC 1157
|
|
|
DB 1081 GGTGATGAGGCTTCATCACTTGAAACCACTGGAACAGATGCCCTACACCACTATGTC 1140
|
|
|
OY 1158 ATTAAGAGGCACTGAGGCTCTACCCACCGGTGCAAGCATTTGCAAGAGCTCAGCACT 1217
|
|
|
DB 1141 ATTAAGAGGCACTGAGGCTCTACCCACCGGTGCAAGCATTTGCAAGAGCTCAGCACT 1200
|
|
|
OY 1218 CCCGTCACCTTCCCTGATGGGCGCTCTTGCCCAAGGTATCATGTCTCTCTCCATT 1277
|
|
|
DB 1201 CCCGTCACCTTCCCTGATGGGCGCTCTTGCCCAAGGTATCATGTCTCTCTCCATT 1260
|
|
|
OY 1278 TATGGCCTTCAACCAACCCAAAGTGTGGCCCAAGAGGTGTTGACCTTTCCGT 1337
|
|
|
DB 1261 TATGGCCTTCAACCAACCCAAAGTGTGGCCCAAGAGGTGTTGACCTTTCCGT 1320
|
|
|
OY 1338 TTTGACACCGGTTCTGCTCAACAACAGCCAGCTTTCTGCTCTCAGAGAGATCAAG 1397
|
|
|
DB 1321 TTTGACACCGGTTCTGCTCAACAACAGCCAGCTTTCTGCTCTCAGAGAGATCAAG 1380
|
|
|
OY 1398 AACTGCAATTGGGAAACAATTGCCATGAACGAGTGAAGTGGCCACGGCCTGACCTG 1457
|
|
|
DB 1381 AACTGCAATTGGGAAACAATTGCCATGAACGAGTGAAGTGGCCACGGCCTGACCTG 1440
|
|
|
OY 1458 CTCCGCTTTGAGCTGCTGCTGATCCCAACAGATCCCATCCCATTTGCAAGACTGTG 1517
|
|
|
DB 1441 CTCCGCTTTGAGCTGCTGCTGATCCCAACAGATCCCATCCCATTTGCAAGACTGTG 1500
|
|
|
OY 1518 TTGAATCCAAAATGAATCCACCTGCTCAAGAGGCTCCCTAACCTTTGTGAAGAC 1577
|
|
|
DB 1501 TTGAATCCAAAATGAATCCACCTGCTCAAGAGGCTCCCTAACCTTTGTGAAGAC 1560
|
|
|
OY 1578 AAGAACCAAGCTTTGAGGGGCTCCACCTGCGTCTGTTCTTCTGACCCCGCTTCTGTC 1637
|
|
|
DB 1561 AAGAACCAAGCTTTGAGGGGCTCCACCTGCGTCTGTTCTTCTGACCCCGCTTCTGTC 1620
|
|
|
OY 1638 CCTTCTGTCTGCCCCATATCTGTTTCTGCTGCCCCACTTCCCTTCTTCCACCTGCG 1697
|
|
|
DB 1621 CCTTCTGTCTGCCCCATATCTGTTTCTGCTGCCCCACTTCCCTTCTTCCACCTGCG 1680
|
|
|
OY 1698 TGCTGTCCCCAGTCTGCTGCTCTCTCTCTCACTTTCTCAGGCTCCCTACCTGTC 1757
|
|
|
DB 1681 TGCTGTCCCCAGTCTGCTGCTCTCTCTCTCACTTTCTCAGGCTCCCTACCTGTC 1740
|
|
|
OY 1758 TTGTCTACCTGTCTCTACCACTGTATCTCTTGTGGGAGAAAAGCTGAGTGTGGGA 1817
|
|
|
DB 1741 TTGTCTACCTGTCTCTACCACTGTATCTCTTGTGGGAGAAAAGCTGAGTGTGGGA 1800
|
|
|
OY 1818 GAAAGTGAAGGCGGAGCTTGCAATGTCTGAATATGTAAAGAGTCTTGAATCATGTCCAG 1877
|
|
|
DB 1801 GAAAGTGAAGGCGGAGCTTGCAATGTCTGAATATGTAAAGAGTCTTGAATCATGTCCAG 1860
|
|
|
```

```
OY 1878 GATCCAGGCTTAACCCCTTGTGCTTTGGAACCAAGCTCTGTGCTGAAGGTGG 1937
|
|
|
DB 1861 GATCCAGGCTTAACCCCTTGTGCTTTGGAACCAAGCTCTGTGCTGAAGGTGG 1920
|
|
|
OY 1938 AAGGCTACCTTGACGCAACCATATCTTAAGCCGGGGCATTAACCCCTCTGCTGGAT 1997
|
|
|
DB 1921 AAGGCTACCTTGACGCAACCATATCTTAAGCCGGGGCATTAACCCCTCTGCTGGAT 1979
|
|
|
OY 1998 AGAATCCAGGGCTCTGCTGCTGGAATGTCTGGAATCTGCTGCTGCTCTCTGCT 2057
|
|
|
DB 1980 AGAATCCAGGGCTCTGCTGCTGGAATGTCTGGAATCTGCTGCTGCTCTGCTCT 2039
|
|
|
OY 2058 CTCCAGGATCAATGTATCTTGAATTAAGAACTGCTCTCCATTAATCTCAAGTACA 2117
|
|
|
DB 2040 CTCCAGGATCAATGTATCTTGAATTAAGAACTGCTCTCCATTAATCTCAAGTACA 2099
|
|
|
OY 2118 GAGCAGATGCTAAGCGTCAACAGCTGTAATGTGTGCTTAATGCAACATGCCCTTTCGA 2177
|
|
|
DB 2100 GAGCAGATGCTAAGCGTCAACAGCTGTAATGTGTGCTTAATGCAACATGCCCTTTCGA 2159
|
|
|
OY 2178 CCCACCCCATTTCTCAACCACTGTTCTTTGTTGATCAACCAATTAATCTGCACTT 2237
|
|
|
DB 2160 C---CCCAATTTCAACCACTGTTCTTTGTTGATCAACCAATTAATCTGCACTT 2216
|
|
|
OY 2238 CCAAGCTCGGGGCTTCAACAGCTCCATCTTAAGCTTTGGCGCCCTGACCCACTTCT 2297
|
|
|
DB 2217 CCAAGCTCGGGGCTTCAACAGCTCCATCTTAAGCTTTGGCGCCCTGACCCACTTCT 2276
|
|
|
OY 2298 CTCTCAACTGCTTTCTCACTGCTTGAATCTGCGGACTTTGTACACCCCAAGACT 2357
|
|
|
DB 2277 CTCTCAACTGCTTTCTCACTGCTTGAATCTGCGGACTTTGTACACCCCAAGACT 2336
|
|
|
OY 2358 GGTGTTGGGTCTGAACACCCCAACATCCCTGAATCTCCACCCACC 2402
|
|
|
DB 2337 GGTGTTGGGTCTGAACACCCCAACATCCCTGAATCTCCACCCACC 2381
|
|
|
```

```
RESULT 7
AX336099 1763 bp DNA linear PAT 09-JAN-2002
LOCUS AX336099
DEFINITION Sequence 6608 from Patent WO0194629.
ACCESSION AX336099
VERSION AX336099.1 GI:18126818
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Hortigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
JOURNAL Patent: WO 0194629-A 6608 13-DEC-2001;
FEATURES
source
1..1763
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 381 a 538 c 430 g 414 t
ORIGIN
Query Match 63.1%; Score 1625.8; DB 6; Length 1763;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;
```

```
OY 10 AGAGATCCAGCAGGTGCTGCAACATGAGTGTCTGTGCTGAGCCCCAGACAGACTCTGG 69
|
|
|
DB 1 AGAGATCCAGCAGGTGCTGCAACATGAGTGTCTGTGCTGAGCCCCAGACAGACTCTGG 60
|
|
|
OY 70 GTGATGTCTTGAATCTTCAAGCGGCTCCCTGCTCAATCTGCTTCTGCTGCTGATCA 129
|
|
|
```

```

Db      61 GTGATGTCCTGGAATCCTCCAGCGGCTCCCTGCTCATTCGTCTGCTGATCA 120
QY      130 AGGCACTTCACTCTAAGCTGACAGGCAAGTGGCTGCTCAAAAGCCCTCCAGCACTTCCCGT 189
Db      121 AGGCACTTCACTCTAAGCTGACAGGCAAGTGGCTGCTCAAAAGCCCTCCAGCACTTCCCGT 180
QY      190 GCCCTCCCTCCCACTGGCTCTTGGGGCAATCCAGAGCTCCAAAGAGCCAGAGCTAC 249
Db      181 GCCCTCCCTCCCACTGGCTCTTGGGGCAATCCAGAGCTCCAAAGAGCCAGAGCTAC 240
QY      250 AACGATTCAGAAATGGGTGAGACATTCCTCAAGTGCCTGCTCAATGGCTATGGGAG 309
Db      241 AACGATTCAGAAATGGGTGAGACATTCCTCAAGTGCCTGCTCAATGGCTATGGGAG 300
QY      310 GCAAGTTGCTGTCAGCTCTATGACCTGACTATATGAAGTGAATTCGGGAGATCAG 369
Db      301 GCAAGTTGCTGTCAGCTCTATGACCTGACTATATGAAGTGAATTCGGGAGATCAG 360
QY      370 ACCGGAATCCCATGGTTCCTACAGATTCCTGGCTCCATGATGGGTACGGCTTGTCC 429
Db      361 ACCGGAATCCCATGGTTCCTACAGATTCCTGGCTCCATGATGGGTACGGCTTGTCC 420
QY      430 TGTGAATGGGAGACATGGTTCAGACATCGAGATGTCGACCCAGCCCTTCCACTATG 489
Db      421 TGTGAATGGGAGACATGGTTCAGACATCGAGATGTCGACCCAGCCCTTCCACTATG 480
QY      490 ACATCCTGAAGCCCTATATGTGGGCTCATGGCAGACTGTGACGAGTGTGACAAAT 549
Db      481 ACATCCTGAAGCCCTATATGTGGGCTCATGGCAGACTGTGACGAGTGTGACAAAT 540
QY      550 GGAAGAGCTCTTGGCCAGGATTCCTCTGAGGTCTTTCAGACGCTCTCTGATGA 609
Db      541 GGAAGAGCTCTTGGCCAGGATTCCTCTGAGGTCTTTCAGACGCTCTCTGATGA 600
QY      610 CCTGGAACCATCATGAAGTGCCTTCAGCCATCAGGGCAGCATCCAGGTGACAGGA 669
Db      601 CCTGGAACCATCATGAAGTGCCTTCAGCCATCAGGGCAGCATCCAGGTGACAGGA 660
QY      670 ATTCTCAGTCTCATATACAGGCCATTAAGTGAAGCTGAACCACTGGTTTTTCCCGTGA 729
Db      661 ATTCTCAGTCTCATATACAGGCCATTAAGTGAAGCTGAACCACTGGTTTTTCCCGTGA 720
QY      730 GGAATGCTTTCACCAAGATGACCACTTACAGCCTGACCTGCTGCGCGCTGACAC 789
Db      721 GGAATGCTTTCACCAAGATGACCACTTACAGCCTGACCTGCTGCGCGCTGACAC 780
QY      790 ACCGCGCTGCAAGCTGGCCCATCAGCAC--AGACCAAGTATCCAACTGAGGAAG 846
Db      781 ACCGCGCTGCAAGCTGGCCCATCAGCACCAAGATGATCCAACTGAGGAAG 840
QY      847 CTCAACTACAGAGGAGGGGAGCTGAGAGATCAAGAGAGAGGCAATTTGATTTTC 906
Db      841 CTCAACTACAGAGGAGGGGAGCTGAGAGATCAAGAGAGAGGCAATTTGATTTTC 900
QY      907 TGGATATCTCTCTTGGCCAAATGAGAAATGGAGCATTTGTCAACAAGACCTCC 966
Db      901 TGGATATCTCTCTTGGCCAAATGAGAAATGGAGCATTTGTCAACAAGACCTCC 960
QY      967 GTGCTGAGGTGACAGCTTCAATGTTGAGGGCCACGACACCAAGCCAGTGGATCTCT 1026
Db      961 GTGCTGAGGTGACAGCTTCAATGTTGAGGGCCACGACACCAAGCCAGTGGATCTCT 1020
QY      1027 GGATCCTTATGCTCTGGCCACACCCCAAGCATCAGAGAGGTGCCGGAGAGATCC 1086
Db      1021 GGATCCTTATGCTCTGGCCACACCCCAAGCATCAGAGAGGTGCCGGAGAGATCC 1080
QY      1087 ACAGCCTCTGGGTGATGAGGCTCCATCACTGGAACCACTGGAACCAAGTCCCTTACA 1146
Db      1081 ACAGCCTCTGGGTGATGAGGCTCCATCACTGGAACCACTGGAACCAAGTCCCTTACA 1140
QY      1147 CCACCATGTCATTAAAGAGGCACTGAGGCTTACCCACCGGTCCAGGCACTTGGCAGAG 1206
Db      1141 CCACCATGTCATTAAAGAGGCACTGAGGCTTACCCACCGGTCCAGGCACTTGGCAGAG 1200

```

```

QY      1207 AGCTCAGACTCCCGTCACTTCCCTGATGGGCGCTCCTTGCCCAAGGATCATGTGCC 1266
Db      1201 AGCTCAGACTCCCGTCACTTCCCTGATGGGCGCTCCTTGCCCAAGGATCATGTGCC 1260
QY      1267 TCCTCTCATTTATGCGCTTCAACCAACCCAAAGTGTGGCCCAACCCAGAGTGTTC 1326
Db      1261 TCCTCTCATTTATGCGCTTCAACCAACCCAAAGTGTGGCCCAACCCAGAGTGTTC 1320
QY      1327 ACCCTTCCGTTTGGACCGGGTCTGCTCAACACAGCCAGCTTTCCTGCTTCTCAG 1386
Db      1321 ACCCTTCCGTTTGGACCGGGTCTGCTCAACACAGCCAGCTTTCCTGCTTCTCAG 1380
QY      1387 GAGATCAAGGAATGCAATGGGAAACAATTTGCCATGAACGAGCTGAAGTGGCC--A 1443
Db      1381 GAGATCAAGGAATGCAATGGGAAACAATTTGCCATGAACGAGCTGAAGTGGCCAGC 1440
QY      1444 CGGCCCTGACCTGCTCCGCTTTGAGCTGCTGATCCCAACAGATCCCATCCCA 1503
Db      1441 AGGCCCTGACCTGCTCCGCTTTGAGCTGCTGATCCCAACAGATCCCATCCCA 1500
QY      1504 TTGCACGACTTGTGTGAATCCAAAAATGGAATCCACCTGCTCTCAGAGGCTCCCTA 1563
Db      1501 TTGCACGACTTGTGTGAATCCAAAAATGGAATCCACCTGCTCTCAGAGGCTCCCTA 1560
QY      1564 ACCCTTGTGAAGCAAGGACCAAGCTTTGAGGCGCTCCACCTGCGCTCTTCTCTGAC 1623
Db      1561 ACCCTTGTGAAGCAAGGACCAAGCTTTGAGGCGCTCCACCTGCGCTCTTCTCTGAC 1620
QY      1624 CCCCGCTTCTGTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1683
Db      1621 CCCCGCTTCTGTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
QY      1684 TCTTC 1688
Db      1681 TCTCC 1685

```

```

RESULT 8
AX336427 1763 bp DNA linear PAT 09-JAN-2002
LOCUS AX336427 6936 from Patent WO0194629.
DEFINITION Sequence
ACCESSION AX336427
VERSION AX336427.1 GI:18127146
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE 1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrihan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 6936 13-DEC-2001;
FEATURES
Source Avalon Pharmaceuticals (US)
Location/Qualifiers
1. 1763
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 381 a 538 c 430 g 414 t
ORIGIN
Query Match 63.1%; Score 1625.8; DB 6; Length 1763;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;
QY 10 AGAGATCAGAGGTGCTGACCATGATGCTCTGTGTGAGGCCAGAGACTCCTGG 69
Db 1 AGAGATCAGAGGTGCTGACCATGATGCTCTGTGTGAGGCCAGAGACTCCTGG 60
QY 70 GTGATGTCCTGGAATCCTCAAGCGGCTCCCTGCTCATTCGTCTGCTGATCA 129

```



Db 61 GTGATGTCCTGGAATCTCCAGCGGCTCCCTGCTCATTCGTCTGCTGATCA 120  
Qy 130 AGGCAAGTTCAGCTCTACCTGCA CAGGCAAGGCTGCTCAAAAGCCCTCAGCATGTTCCCGT 189  
Db 121 AGGCAAGTTCAGCTCTACCTGCA CAGGCAAGGCTGCTCAAAAGCCCTCAGCATGTTCCCGT 180  
Qy 190 GCCCTCCCTCCCACTGGCTCTTCGGGCA CATCCAGAGGCTCCCAAGACCAAGAGCTAC 249  
Db 181 GCCCTCCCTCCCACTGGCTCTTCGGGCA CATCCAGAGGCTCCCAAGACCAAGAGCTAC 240  
Qy 250 AACGGAATCAGAAATGGGTGAGACATTC CCAAGTGCCTGTCTCATTTGGCTATGGGAG 309  
Db 241 AACGGAATCAGAAATGGGTGAGACATTC CCAAGTGCCTGTCTCATTTGGCTATGGGAG 300  
Qy 310 GCAAAGTTCGTTCAGCTCTATGACCTG AACTATATGAAGGTGATCTGGGAGATCAG 369  
Db 301 GCAAAGTTCGTTCAGCTCTATGACCTG AACTATATGAAGGTGATCTGGGAGATCAG 360  
Qy 370 ACCGGAATCCCATGGTTCCTACAGATT CCTGGCTCCATGATGGGTACGGCTTGCTCC 429  
Db 361 ACCGGAATCCCATGGTTCCTACAGATT CCTGGCTCCATGATGGGTACGGCTTGCTCC 420  
Qy 430 TGTGAATGGGCA CAGATGTTCCAGCAT CAGCGATGCTGACCCCAAGCTTCCACTATG 489  
Db 421 TGTGAATGGGCA CAGATGTTCCAGCAT CAGCGATGCTGACCCCAAGCTTCCACTATG 480  
Qy 490 ACATCCTGAAGCCCTATGTGGGCTCAT GGGAGACACTGTACAGATGCTGACAAAT 549  
Db 481 ACATCCTGAAGCCCTATGTGGGCTCAT GGGAGACACTGTACAGATGCTGACAAAT 540  
Qy 550 GGAAGAGCTCCTGGCCAGGATCCCTCT GAGGCTCTTCAAGACGTCTCCTGATGA 609  
Db 541 GGAAGAGCTCCTGGCCAGGATCCCTCT GAGGAGGTCTTCAAGACGTCTCCTGATGA 600  
Qy 610 CCCTGGA CACCATGAGTGTGCTTCA GGCATCAAGGCAAGCATCAAGTGA CAGGA 669  
Db 601 CCCTGGA CACCATGAGTGTGCTTCA GGCATCAAGGCAAGCATCAAGTGA CAGGA 660  
Qy 670 ATTCTCAGTCTTACATACAGGCCATT AGTGACCTGAACCAACCTGTTTTTCCCGTGTGA 729  
Db 661 ATTCTCAGTCTTACATACAGGCCATT AGTGACCTGAACCAACCTGTTTTTCCCGTGTGA 720  
Qy 730 GGAATGCTTTACAGGAATGACACCAT CTACAGCTGACCTGTGCTGGCCGCTGACAC 789  
Db 721 GGAATGCTTTACAGGAATGACACCAT CTACAGCTGACCTGTGCTGGCCGCTGACAC 780  
Qy 790 ACCGGCCTG CCACTGGCCCATCAGACA C---AGACCAAGTATCCAACTGAGGAAG 846  
Db 781 ACCGGCCTG CCACTGGCCCATCAGACA CAGGACCAAGTATCCAACTGAGGAAG 840  
Qy 847 CTCAACTACAGAAAGAGGGGAGCTGGA GAAGATCAAGAAGAGGCAATTTGGATTTC 906  
Db 841 CTCAACTACAGAAAGAGGGGAGCTGGA GAAGATCAAGAAGAGGCAATTTGGATTTC 900  
Qy 907 TGGATATCTCTCTTGGCCAAATGGAAT GGAAGCATTTGTCAACAAGGACCTCC 966  
Db 901 TGGATATCTCTCTTGGCCAAATGGAAT GGAAGCATTTGTCAACAAGGACCTCC 960  
Qy 967 GTGCTGAGTGA CACGTTCATGTTGAGG GGCACGACACCAAGGAGTGGGATCTCCT 1026  
Db 961 GTGCTGAGTGA CACGTTCATGTTGAGG GGCACGACACCAAGGAGTGGGATCTCCT 1020  
Qy 1027 GGATCCTTATG CTGGCCACACACCCCA AGCATCAGAGAGGTGCCGGAAGAGATCC 1086  
Db 1021 GGATCCTTATG CTGGCCACACACCCCA AGCATCAGAGAGGTGCCGGAAGAGATCC 1080  
Qy 1087 ACAAGCTCTGG GTAGTGAAGCTCCATCA CCTGGAACCACTGACCAAGTCCCTACA 1146  
Db 1081 ACAAGCTCTGG GTAGTGAAGCTCCATCA CCTGGAACCACTGACCAAGTCCCTACA 1140  
Qy 1147 CCACCATGTGA TTAAGAGGCACTGAGG CTCTAACCAACCGGTGCCAGGCAATTTGGCAGAG 1206

Db 1141 CCACCATGTGATTAAGAGGCACTGAG GCTCTAACCCACCGGTGCCAGGATGGCAGAG 1200  
Qy 1207 AGCTCAGACATCCCGTCACTTCCCTGA TGGGCGCTCCTTGCCCAAGATATCATGTCC 1266  
Db 1201 AGCTCAGACATCCCGTCACTTCCCTGA TGGGCGCTCCTTGCCCAAGATATCATGTCC 1260  
Qy 1267 TCCTCTCCATTATGGCCTTCAACCAACA AAGTGTGGGCCCAAGAGGTGTTG 1326  
Db 1261 TCCTCTCCATTATGGCCTTCAACCAACA AAGTGTGGGCCCAAGAGGTGTTG 1320  
Qy 1327 ACCCTTCCGTTTGCACCGGGTCTGTCA ACAAGCCAGCTTTCTGCTCTCAG 1386  
Db 1321 ACCCTTCCGTTTGCACCGGGTCTGTCA ACAAGCCAGCTTTCTGCTCTCAG 1380  
Qy 1387 GAGGATCAAGAACTGCATTGGGAAACA ATTGACATGACAGAGTGAAGTGCC--A 1443  
Db 1381 GAGGATCAAGAACTGCATTGGGAAACA ATTGACATGACAGAGTGAAGTGCCAGC 1440  
Qy 1444 CGGCCCTGACCCCTGCTCCGCTTGAGCT GCTGCTGATCCCAAGAGATCCCATCCCA 1503  
Db 1441 AGGCCCTGACCCCTGCTCCGCTTGAGCT GCTGCTGATCCCAAGAGATCCCATCCCA 1500  
Qy 1504 TTGCACGACTTGTGTGAATCCAAAATGGA ATCCACCTGCGTCTCAGAGGCTCCCTA 1563  
Db 1501 TTGCACGACTTGTGTGAATCCAAAATGGA ATCCACCTGCGTCTCAGAGGCTCCCTA 1560  
Qy 1564 ACCCTTGTGAAGACAAAGACCAAGCTT TGAAGGCTCCACCTGCGCTGCTCTGAC 1623  
Db 1561 ACCCTTGTGAAGACAAAGACCAAGCTT TGAAGGCTCCACCTGCGCTGCTCTGAC 1620  
Qy 1624 CCCCCTCTGTGCTCCCTCTCCTGTCTG CCGCATATCTGTTTCTGTGCTGCCCACTTCCCT 1683  
Db 1621 CCCCCTCTGTGCTCCCTCTCCTGTCTG CCGCATATCTGTTTCTGTGCTGCCCACTTCCCTG 1680  
Qy 1684 TCTTC 1688  
Db 1681 TCTTC 1685

RESULT 9  
AX336709 1763 bp DNA linear PAT 09-JAN-2002  
LOCUS AX336709 Sequence 7218 from Patent WO0194629.  
DEFINITION AX336709  
ACCESSION AX336709  
VERSION AX336709.1 GI:18127428  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
REFERENCE  
AUTHORS Young, P.B., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,  
Horrihan, S., Soppet, D.R. and Weaver, Z.  
TITLE Cancer gene determination and therapeutic screening using signature  
JOURNAL Patent: WO 0194629-A 7218 13-DEC-2001;  
FEATURES  
source  
1. .1763  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 381 a 538 c 430 g 414 t  
ORIGIN  
Query Match 63.1%; Score 1625.8; DB 6; Length 1763;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;  
Qy 10 AGAGATCCAGCAGGTGCTGACCATGAGTGTCTGTGCTGAGCCCCAGCAGACTCCTGG 69  
Db 1 AGAGATCCAGCAGGTGCTGACCATGAGTGTCTGTGCTGAGCCCCAGCAGACTCCTGG 60

Qy 70 GTGATGTCCTGGAATCTCCAGCGGCTCCCTGCTCATTCCTGCTGCTGATCA 129  
Db 61 GTGATGTCCTGGAATCTCCAGCGGCTCCCTGCTCATTCCTGCTGCTGATCA 120  
Qy 130 AGGAGTTCAGCTCTACCTGCA CAGGAGTGCTGCTCAAGCCCTCCAGAGTTCCCT 189  
Db 121 AGGAGTTCAGCTCTACCTGCA CAGGAGTGCTGCTCAAGCCCTCCAGAGTTCCCT 180  
Qy 190 GCCCTCCCTCCCACTGCTCTTCGGGACATCCAGAGCTCCACAGAGCCAGAGCTAC 249  
Db 181 GCCCTCCCTCCCACTGCTCTTCGGGACATCCAGAGCTCCACAGAGCCAGAGCTAC 240  
Qy 250 AACGATTCAGAAATGGGTGAGACATTCCTCAAGTGCCTGCTCATTTGGCTATGGGAG 309  
Db 241 AACGATTCAGAAATGGGTGAGACATTCCTCAAGTGCCTGCTCATTTGGCTATGGGAG 300  
Qy 310 GCAAAGTTCGTGTCAGCTCTATGACCTGACTATATGAAGTGATTCGGGAGATCAG 369  
Db 301 GCAAAGTTCGTGTCAGCTCTATGACCTGACTATATGAAGTGATTCGGGAGATCAG 360  
Qy 370 ACCCGAAATCCCATGTTCTCTACAGATTCCTGCTCCATGATGGGTACGGCTTCTCC 429  
Db 361 ACCCGAAATCCCATGTTCTCTACAGATTCCTGCTCCATGATGGGTACGGCTTCTCC 420  
Qy 430 TGTGAAATGGGACAGCATGTTCCAGATCGACGATGCTGACCCAGCTTCCACTATG 489  
Db 421 TGTGAAATGGGACAGCATGTTCCAGATCGACGATGCTGACCCAGCTTCCACTATG 480  
Qy 490 ACATCTGAAAGCCCTATGTTGGGCTCATGAGCACTCTGACAGATGCTGACAAAT 549  
Db 481 ACATCTGAAAGCCCTATGTTGGGCTCATGAGCACTCTGACAGATGCTGACAAAT 540  
Qy 550 GGGAGAGCTCTTGCCAGAGATCCCTCTGAGGTCTTTCAGCACTCTCTGATGA 609  
Db 541 GGGAGAGCTCTTGCCAGAGATCCCTCTGAGGTCTTTCAGCACTCTCTGATGA 600  
Qy 610 CCCTGACACCATCATGAGTGTCTTCAAGCCATCAGGGGACGATCCAGGTGACAGA 669  
Db 601 CCCTGACACCATCATGAGTGTCTTCAAGCCATCAGGGGACGATCCAGGTGACAGA 660  
Qy 670 ATTCTAGTCTCTACATACAGGCCATTAAGTGAACCTGAACCTGTTTTCCTGCTGA 729  
Db 661 ATTCTAGTCTCTACATACAGGCCATTAAGTGAACCTGAACCTGTTTTCCTGCTGA 720  
Qy 730 GGAATGCTTTTCA CAGAGATGACCATCTACAGCTGACCTTCTGTCGCGCTGACAC 789  
Db 721 GGAATGCTTTTCA CAGAGATGACCATCTACAGCTGACCTTCTGTCGCGCTGACAC 780  
Qy 790 ACCGCGCTGCCAGCTGGCCCATCAGACAC--AGACCAAGTATCCAACTGAGAAAG 846  
Db 781 ACCGCGCTGCCAGCTGGCCCATCAGACAC--AGACCAAGTATCCAACTGAGAAAG 840  
Qy 847 CTCAACTACAGAGAGAGGGGAGCTGAGAGAGATCAAGAGAGAGGCAATTGATTTTC 906  
Db 841 CTCAACTACAGAGAGAGGGGAGCTGAGAGAGATCAAGAGAGAGGCAATTGATTTTC 900  
Qy 907 TGGATATCTCTCTTGGCCAAATGAGAGATGGAGCATCTGTCAAGCAAGACCTCC 966  
Db 901 TGGATATCTCTCTTGGCCAAATGAGAGATGGAGCATCTGTCAAGCAAGACCTCC 960  
Qy 967 GTGCTGAGGTGACACGTTTATGTTGAGGGCCACGACACCAAGCCAGTGGATCTCT 1026  
Db 961 GTGCTGAGGTGACACGTTTATGTTGAGGGCCACGACACCAAGCCAGTGGATCTCT 1020  
Qy 1027 GGATCTCTATATGCTTGCCACACACCCCAAGCATCAGAGAGAGTCCCGGAGAGATCC 1086  
Db 1021 GGATCTCTATATGCTTGCCACACACCCCAAGCATCAGAGAGAGTCCCGGAGAGATCC 1080  
Qy 1087 ACAGCTCTCTGGGTGATGAGCTTCATCACTGGAACCACTGGAACAGATGCCCTACA 1146  
Db 1081 ACAGCTCTCTGGGTGATGAGCTTCATCACTGGAACCACTGGAACAGATGCCCTACA 1140  
Qy 1147 CCACCATGTCATTAAAGAGGCACTGAGGCTTACCAACCGGTGCCAGGCAATTGGCAGAG 1206

Db 1141 CCACCATGTCATTAAAGAGGCACTGAGGCTCTAACCAACCGGTGCCAGAGATTGGCAGAG 1200  
Qy 1207 AGCTCAGCACTCCGTCACCTTCCCTGATGGGCGCTCCTTGCCCAAAGTATCATGTCC 1266  
Db 1201 AGCTCAGCACTCCGTCACCTTCCCTGATGGGCGCTCCTTGCCCAAAGTATCATGTCC 1260  
Qy 1267 TCCTCTCATTTATGGCCTTACCAACAACCAAAAGTGTGSCCAACCAAGAGTGTTC 1326  
Db 1261 TCCTCTCATTTATGGCCTTACCAACAACCAAAAGTGTGSCCAACCAAGAGTGTTC 1320  
Qy 1327 ACCCTTCCGTTTTCACACCGGGTCTGCTCAACAACAGCCAGCTTCTGCTCTCTCAG 1386  
Db 1321 ACCCTTCCGTTTTCACACCGGGTCTGCTCAACAACAGCCAGCTTCTGCTCTCTCAG 1380  
Qy 1387 GAGATCAAGGAACCTGCAATTGGGAACAATTGCCATGAACGAGCTGAAGGTGGCC--A 1443  
Db 1381 GAGATCAAGGAACCTGCAATTGGGAACAATTGCCATGAACGAGCTGAAGGTGGCCAGC 1440  
Qy 1444 CGGCCCTGACCTGCTCCGCTTGAAGCTGCTGCTGATCCCAACAGATCCCATCCCA 1503  
Db 1441 AGGCCCTGACCTGCTCCGCTTGAAGCTGCTGCTGATCCCAACAGATCCCATCCCA 1500  
Qy 1504 TTGCACGACTGTGTGAATCCAAATGGAATCCACCTGCGTCTCAGAGGCTCCCTA 1563  
Db 1501 TTGCACGACTGTGTGAATCCAAATGGAATCCACCTGCGTCTCAGAGGCTCCCTA 1560  
Qy 1564 ACCCTGTGAAGACAAGGACCAAGCTTTGAGGGCTCCACCTGCGCTGTCTCTGAC 1623  
Db 1561 ACCCTGTGAAGACAAGGACCAAGCTTTGAGGGCTCCACCTGCGCTGTCTCTGAC 1620  
Qy 1624 CCCGCTTCTGTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1683  
Db 1621 CCCGCTTCTGTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680  
Qy 1684 TCTTC 1688  
Db 1681 TCTCC 1685

RESULT 10  
AX408963 1763 bp DNA linear PAT 14-JUN-2002  
LOCUS AX408963 1610 from Patent WO0229103.  
DEFINITION Sequence  
ACCESSION AX408963  
VERSION AX408963 1 GI:21441668  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1  
AUTHORS Alvares, C., Horne, D., Perez-da-Silva, S. and Vockley, J. G.  
TITLE Gene expression profiles in liver cancer  
JOURNAL Patent: WO 0229103-A 1610 11-APR-2002;  
GENE LOGIC INC (US)  
FEATURES  
source Location/Qualifiers  
1. 1763  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/note="EMBL/GenBank Accession No. D13705"  
BASE COUNT 381 a 538 c 430 g 414 t  
ORIGIN  
Query Match 63.1%; Score 1625.8; DB 6; Length 1763;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;  
Qy 10 AGAGATCGACGAGTGCTGACCATGATGTTCTGTGTGAGCCCAAGACTCTGG 69  
Db 1 AGAGATCGACGAGTGCTGACCATGATGTTCTGTGTGAGCCCAAGACTCTGG 60

OY 70 GTGATGTCTGGAATCTCCCAAGCGGCTCCCTGCTCATTTCTGCTTCTGCTGATCA 129  
|||||  
DB 61 GTGATGTCTGGAATCTCCCAAGCGGCTCCCTGCTCATTTCTGCTTCTGCTGATCA 120  
OY 130 AGGCAAGTCAAGCTCTAAGCTGCAAGGCAAGTGGCTCTCAAGCCCTCCAGCAAGTTCCCGT 189  
|||||  
DB 121 AGGCAAGTCAAGCTCTAAGCTGCAAGGCAAGTGGCTCTCAAGCCCTCCAGCAAGTTCCCGT 180  
OY 190 GCCCTCCCTCCCAAGTGGCTCTTGGGCAATCCAGAGCTCCCAAGAGCCAGAGCTTAC 249  
|||||  
DB 181 GCCCTCCCTCCCAAGTGGCTCTTGGGCAATCCAGAGCTCCCAAGAGCCAGAGCTTAC 240  
OY 250 AACGATTCAGAAATGGGTGAGAGACATTCGCAAGTGCCTGCTCATGCTATGGGAG 309  
|||||  
DB 241 AACGATTCAGAAATGGGTGAGAGACATTCGCAAGTGCCTGCTCATGCTATGGGAG 300  
OY 310 GCAAGTTCGTCTCAAGCTCTATGACCCCTGACTATATGAAGGTGATTCGGGAGATCAG 369  
|||||  
DB 301 GCAAGTTCGTCTCAAGCTCTATGACCCCTGACTATATGAAGGTGATTCGGGAGATCAG 360  
OY 370 ACCCGAAATCCCATGCTTCTACAGATTCTGCGCTCATGATGGGTACGGCTTGCTCC 429  
|||||  
DB 361 ACCCGAAATCCCATGCTTCTACAGATTCTGCGCTCATGATGGGTACGGCTTGCTCC 420  
OY 430 TGTGAATGGGCAAGCATGGTTCAGCATCGAGGATGCTGACCCAGCTTCCACTATG 489  
|||||  
DB 421 TGTGAATGGGCAAGCATGGTTCAGCATCGAGGATGCTGACCCAGCTTCCACTATG 480  
OY 490 ACATCTGAAGCCCTATGTGGGGCTCATGGGCAAGCTCTGTACGAGTGAATGCTGACAAT 549  
|||||  
DB 481 ACATCTGAAGCCCTATGTGGGGCTCATGGGCAAGCTCTGTACGAGTGAATGCTGACAAT 540  
OY 550 GGGAGAGCTCTTGGGCAAGGATTCCTCTGAGAGGTCTTCAAGCAGTCTCTTGATGA 609  
|||||  
DB 541 GGGAGAGCTCTTGGGCAAGGATTCCTCTGAGAGGTCTTCAAGCAGTCTCTTGATGA 600  
OY 610 CCCTGGAACCATCATGAGTGTGCTTCAAGCATGAGGCAAGCATCCAGTGAACAGGA 669  
|||||  
DB 601 CCCTGGAACCATCATGAGTGTGCTTCAAGCATGAGGCAAGCATCCAGTGAACAGGA 660  
OY 670 ATTCTCAGTCTCAATACAGGCAATTAAGTGAACCTGATGAACCACTGTTTCCCGTGA 729  
|||||  
DB 661 ATTCTCAGTCTCAATACAGGCAATTAAGTGAACCTGATGAACCACTGTTTCCCGTGA 720  
OY 730 GGAATGCTTTCAAGCAAGATGACACCATCTACAGCTGACCTCTGCTGGCGCTGACAC 789  
|||||  
DB 721 GGAATGCTTTCAAGCAAGATGACACCATCTACAGCTGACCTCTGCTGGCGCTGACAC 780  
OY 790 ACCGGCGCTGCAAGTGGCCCATCAGACAC--AGACCAAGTGAATCCAACTGAGGAAG 846  
|||||  
DB 781 ACCGGCGCTGCAAGTGGCCCATCAGACAC--AGACCAAGTGAATCCAACTGAGGAAG 840  
OY 847 CTCAACTACAGAGAGGGGAGCTGAGAGATCAAGAGAGAGGCAATTTGATTTC 906  
|||||  
DB 841 CTCAACTACAGAGAGGGGAGCTGAGAGATCAAGAGAGAGGCAATTTGATTTC 900  
OY 907 TGGATATCTCTCTTGGCCAAATGAGATGGAGCATTTGTCAAGACAGACCTCC 966  
|||||  
DB 901 TGGATATCTCTCTTGGCCAAATGAGATGGAGCATTTGTCAAGACAGACCTCC 960  
OY 967 GTGCTGAGGTGACACGTTTATGTTTGAAGGCGCAAGACACAGCCAGTGGGATCTCCT 1026  
|||||  
DB 961 GTGCTGAGGTGACACGTTTATGTTTGAAGGCGCAAGACACAGCCAGTGGGATCTCCT 1020  
OY 1027 GGATCTCTATGCTTGGCCACACACCCCAAGCATCAAGAGAGGTGCGGAGAGATCC 1086  
|||||  
DB 1021 GGATCTCTATGCTTGGCCACACACCCCAAGCATCAAGAGAGGTGCGGAGAGATCC 1080  
OY 1087 ACAGCCTCTGGGTGATGAGCCTTCATCACTGGAACCACTGACCAAGATGCCCTACA 1146  
|||||  
DB 1081 ACAGCCTCTGGGTGATGAGCCTTCATCACTGGAACCACTGACCAAGATGCCCTACA 1140  
OY 1147 CCACCATGTGATTAAGAGGCACTGAGGCTTACCAACCGGTGCCAGGCAATGGCAGAG 1206

DB 1141 CCACATGTGATTAAGAGGCACTGAGGCTTACCAACCGGTGCCAGCATGGCAGAG 1200  
OY 1207 AGCTCAGCACTCCCGTCAACCTTCCCTGATGGGCGCTCTTGCCCAAGGATCATGCTCC 1266  
|||||  
DB 1201 AGCTCAGCACTCCCGTCAACCTTCCCTGATGGGCGCTCTTGCCCAAGGATCATGCTCC 1260  
OY 1267 TCCTCTCATTTATGAGCTTCAACCAACCAAGTGTGGCCCAACCAAGAGTGTTC 1326  
|||||  
DB 1261 TCCTCTCATTTATGAGCTTCAACCAACCAAGTGTGGCCCAACCAAGAGTGTTC 1320  
OY 1327 ACCCTTCCGTTTTCACCCGGGTCTGCTCAACACAGCCAGCTTTCCTGCCCTCTCAG 1386  
|||||  
DB 1321 ACCCTTCCGTTTTCACCCGGGTCTGCTCAACACAGCCAGCTTTCCTGCCCTCTCAG 1380  
OY 1387 GAGGATCAAGAACTGATTTGGGAAACAATTTGCCATGAACGAGCTGAAGGTGCC--A 1443  
|||||  
DB 1381 GAGGATCAAGAACTGATTTGGGAAACAATTTGCCATGAACGAGCTGAAGGTGCCAGC 1440  
OY 1444 CGGCCCTGACCCCTGCTCCGCTTGAAGCTGCTGATCCACAGATCCCATCCCA 1503  
|||||  
DB 1441 AGGCCCTGACCCCTGCTCCGCTTGAAGCTGCTGATCCACAGATCCCATCCCA 1500  
OY 1504 TTGCAAGACTTGTGTAATCCAAATGGAATCCACTGCGTCTCAGAGGCTCCCTA 1563  
|||||  
DB 1501 TTGCAAGACTTGTGTAATCCAAATGGAATCCACTGCGTCTCAGAGGCTCCCTA 1560  
OY 1564 ACCCTTGTGAAGACAAGACCAAGCTTTGAGGGCTCCACCTGCCCTCTCTCTGAC 1623  
|||||  
DB 1561 ACCCTTGTGAAGACAAGACCAAGCTTTGAGGGCTCCACCTGCCCTCTCTCTGAC 1620  
OY 1624 CCCCCTTGTGCTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1683  
|||||  
DB 1621 CCCCCTTGTGCTCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680  
OY 1684 TCTTC 1688  
DB 1681 TCTTC 1685

## RESULT 11

HUMOMHY 1763 bp mRNA linear PRI 03-FEB-1999  
LOCUS  
DEFINITION Human mRNA for fatty acids omega-hydroxylase (cytochrome P-450HKV),  
complete cds.

ACCESSION D13705  
VERSION D13705.1 GI:1197806  
KEYWORDS fatty acid omega-hydroxylase; cytochrome P-450HKV.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS 1 (bases 1 to 1763)  
JOURNAL Kikuta, Y.  
REFERENCE Unpublished  
AUTHORS 2 (bases 1 to 1763)  
TITLE Kikuta, Y.  
JOURNAL Direct Submission  
COMMENT Submitted (18-NOV-1992) Yasushi Kikuta, Fukuyama University,  
Faculty of Engineering; Gakuenchou 1, Fukuyama, Hiroshima 729-0292,  
Japan (Tel:0849-36-2111, Fax:0849-36-2023)  
On Feb 20, 1996 this sequence version replaced gi:535786.  
Submitted (18-NOV-1992) to DDBJ by:  
Yasushi Kikuta  
Department of Food Science and Technology  
Fukuyama University  
Faculty of Engineering  
1 Gakuenchou  
Fukuyama, Hiroshima 729-02  
Japan  
Phone: 0849-36-2111  
Fax: 0849-36-2023  
Location/Qualifiers

## FEATURES

```
source
1. 1763
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HK-7"
/issue_type="kidney"
/dev_stage="adult"
24. 1589
/notes="heme binding cysteine at AA458."
/codon_start=1
/product="fatty acid omega-hydroxylase"
/protein_id="BAA02864.1"
/db_xref="GI:535787"
/translation="MSVSVLSPSRLLGDSIQLAASILLILLIKAVOLYLHROWL
LKALQFPCCPSHWLFGHQLQDQBLQRIQKWEVTPSACPHWLMGKRVQLYDP
DYMKVILGRSDPKSHGYRFLAIVGYGLLLNGQTFQHRMLTPAFHYDLKPYVG
LMADSVRLMDKWBELLGDSPLFVFOHVSMTLDTIMKCAFSHQSIQVDNSOSYI
QAIISDNLNLYESRVNAFHNDITYSLTSAGRWTHRACQLSHQTRPKVIOLRKQLO
KEGLEKIKRRLDPLDILLAKNENSLSDKDLRAVDTFMEGHDTTASISWI
LYALATHPKQBRCBBIHSLDGASITWNHLDQMPYTTMCIKBALRYPVPGIGR
ELSTVTPDGRSLPKGIMVLSIYGLHNPKWPNPEVDFRPAFGSAQSHAPLP
PSGSRNCIGKQFAMNBLKVGGQALTIVRPFLPDPTRIPIPARLVLSKNGIHLRL
RLPNCEDXDL"
polya_signal 1736. 1741
polya_site 1763
BASE COUNT 381 a 538 c 430 g 414 t
ORIGIN
Query Match 63.1%; Score 1625.8; DB 9; Length 1763;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;
QY 10 AGAGATCCAGCAGGTGCTGCACCATGAGTGTCTGTGTGAGCCCGCAGAGACTCCTGG 69
DB 1 AGAGATCCAGCAGGTGCTGCACCATGAGTGTCTGTGTGAGCCCGCAGAGACTCCTGG 60
QY 70 GTGATGTCTGGAATCTCCAAGCGGCTCCTGCTCATTCGTCTGTGCTGTATCA 129
DB 61 GTGATGTCTGGAATCTCCAAGCGGCTCCTGCTCATTCGTCTGTGCTGTATCA 120
QY 130 AGGAGTTGAGCTCTACCTGCACAGGCGAGTGGCTCAAGCCCTCCAGAGTTCCCGT 189
DB 121 AGGAGTTGAGCTCTACCTGCACAGGCGAGTGGCTCAAGCCCTCCAGAGTTCCCGT 180
QY 190 GCCCTCCCTCCCACTGGCTCTTGGGCAATCCAGAGCTCCACAGACCAAGAGCTAC 249
DB 181 GCCCTCCCTCCCACTGGCTCTTGGGCAATCCAGAGCTCCACAGACCAAGAGCTAC 240
QY 250 AACGAGTTCAGAATGGGTGAGACATTCCAAGTGCCTGTCTCATTTGGCTATGGGAG 309
DB 241 AACGAGTTCAGAATGGGTGAGACATTCCAAGTGCCTGTCTCATTTGGCTATGGGAG 300
QY 310 GCAAGTTCGTCTCAGCTCTATGACCCCTGACTATATGAAGTGATCTGGGAGATCAG 369
DB 301 GCAAGTTCGTCTCAGCTCTATGACCCCTGACTATATGAAGTGATCTGGGAGATCAG 360
QY 370 ACCCGAATCCCATGCTTCTACAGATTCTGGCTCCATGATTTGGGTACGGCTTGTCC 429
DB 361 ACCCGAATCCCATGCTTCTACAGATTCTGGCTCCATGATTTGGGTACGGCTTGTCC 420
QY 430 TGTGAATGGGAGACATGTTCCAGCATGACGAGTGTGACCCCGACTTCCACTATG 489
DB 421 TGTGAATGGGAGACATGTTCCAGCATGACGAGTGTGACCCCGACTTCCACTATG 480
QY 490 ACATCCTGAAGCCCTATGTGGGCTCATGGCAGACTCTGTACGAGTATGCTGACAAAT 549
DB 481 ACATCCTGAAGCCCTATGTGGGCTCATGGCAGACTCTGTACGAGTATGCTGACAAAT 540
QY 550 GGAAGAGCTCTTGGCAGAGATCCCTCTGAGAGTCTTTCAGCAGTCTCCTTGATGA 609
DB 541 GGAAGAGCTCTTGGCAGAGATCCCTCTGAGAGTCTTTCAGCAGTCTCCTTGATGA 600
QY 610 CCTGGAACCATGAAAGTGTGCTTCAAGCATCAGGGCAGCATCCAGGTGACAGGA 669
```

```
DB 601 CCCTGGACACCATGATGAAGTGTGCTTCAAGCATCAGGGCAGCATCCAGGTGACAGGA 660
QY 670 ATTCTAGTCCCTACATACAGGCCATTAGTACCTGAACAACCTGGTTTTTCCCGTGTGA 729
DB 661 ATTCTAGTCCCTACATACAGGCCATTAGTACCTGAACAACCTGGTTTTTCCCGTGTGA 720
QY 730 GGAATGCCCTTTCACAGAAATGACACCATCTACAGCCTGACCTGTGCGCGCTGACAC 789
DB 721 GGAATGCCCTTTCACAGAAATGACACCATCTACAGCCTGACCTGTGCGCGCTGACAC 780
QY 790 ACCGCGCTGCACTGTGCCCATCAGCAC--AGACCAAGTATCCAAGAGAGG 846
DB 781 ACCGCGCTGCACTGTGCCCATCAGCACAGAGACCAAGTATCCAAGAGAGG 840
QY 847 CTCAACTACAGAGAGGGAGCTGGAAGATCAAGAGAGAGGATTTGATTTTC 906
DB 841 CTCAACTACAGAGAGGGAGCTGGAAGATCAAGAGAGAGGATTTGATTTTC 900
QY 907 TGGATATCCTCTCTTGGCCAAATGGAATGGAGCATCTTGTACACAAGACCTCC 966
DB 901 TGGATATCCTCTCTTGGCCAAATGGAATGGAGCATCTTGTACACAAGACCTCC 960
QY 967 GTGCTGAGGTGACACGTTTATGTTGAGGGCCACGACACCAAGCTGGGATCTCCT 1026
DB 961 GTGCTGAGGTGACACGTTTATGTTGAGGGCCACGACACCAAGCTGGGATCTCCT 1020
QY 1027 GGATCCTCTATGCTCTGGCCACACACCCCAAGCATCAGAGAGGTGCGGAGAGATCC 1086
DB 1021 GGATCCTCTATGCTCTGGCCACACACCCCAAGCATCAGAGAGGTGCGGAGAGATCC 1080
QY 1087 ACAGCCTCCTGGGTGATGAGCCTCCATCACTGGAACCACTGACCAAGTCCCTACA 1146
DB 1081 ACAGCCTCCTGGGTGATGAGCCTCCATCACTGGAACCACTGACCAAGTCCCTACA 1140
QY 1147 CCACCATGTGATTAAAGAGGACACTGAGGCTTACCCACCGGTGCGAGCATTTGGCAGAG 1206
DB 1141 CCACCATGTGATTAAAGAGGACACTGAGGCTTACCCACCGGTGCGAGCATTTGGCAGAG 1200
QY 1207 AGCTAGACATCCCGTCACTTCCCTGATGAGCGCTCCTTGCCCAAGTATCATGTGCC 1266
DB 1201 AGCTAGACATCCCGTCACTTCCCTGATGAGCGCTCCTTGCCCAAGTATCATGTGCC 1260
QY 1267 TCCTCTCATTTAAGGCTTCAACCAACCAAAAGTGTGCCCCAAGAGGTGTTG 1326
DB 1261 TCCTCTCATTTAAGGCTTCAACCAACCAAAAGTGTGCCCCAAGAGGTGTTG 1320
QY 1327 ACCCTTCCGTTTGCACCGGGTCTGCTCAACACAGCAGCTTCTGCTTCTCAG 1386
DB 1321 ACCCTTCCGTTTGCACCGGGTCTGCTCAACACAGCAGCTTCTGCTTCTCAG 1380
QY 1387 GAGGATCAAGAACTGCAATTGGGAAACAATTGCCATGAACGAGCTGAAGTGGCC--A 1443
DB 1381 GAGGATCAAGAACTGCAATTGGGAAACAATTGCCATGAACGAGCTGAAGTGGCCAGC 1440
QY 1444 CGGCCCTGACCTGCTCGGCTTGAAGCTGCTGATCCCAACAGATCCCATCCCA 1503
DB 1441 AGGCCCTGACCTGCTCGGCTTGAAGCTGCTGATCCCAACAGATCCCATCCCA 1500
QY 1504 TTGCACGACTTGTGTTGAATCCAAAATGGAATCAACTGCGTCTCAGAGAGGCTCCCTA 1563
DB 1501 TTGCACGACTTGTGTTGAATCCAAAATGGAATCAACTGCGTCTCAGAGAGGCTCCCTA 1560
QY 1564 ACCCTGTGAAGACAAGACCAAGCTTTGAGGGCTCCACCTGCGCTGCTCTGAC 1623
DB 1561 ACCCTGTGAAGACAAGACCAAGCTTTGAGGGCTCCACCTGCGCTGCTCTGAC 1620
QY 1624 CCCCCTTCTGTCCCTTCTCTGTCTGCGCATATCTGTGTTCTGTGCGCACCTTCCCT 1683
DB 1621 CCCCCTTCTGTCCCTTCTCTGTCTGCGCATATCTGTGTTCTGTGCGCATACACCTG 1680
QY 1684 TCTTC 1688
```





QY 1147 CCACCATGTGATTAAAGAGGCACTGAGGCTCTAACCCACCGGTGCCAGGCAATTGGCAGAG 1206  
|||||  
DB 1146 CCACCATGTGATTAAAGAGGCACTGAGGCTCTAACCCACCGGTGCCAGGCAATTGGCAGAG 1205  
QY 1207 AGCTCAGCACTCCCGTCACTCCCTGATGGGCGCTCTTGCCCAAGGTATCATGTGCC 1266  
|||||  
DB 1206 AGCTCAGCACTCCCGTCACTCCCTGATGGGCGCTCTTGCCCAAGGTATCATGTGCC 1265  
QY 1267 TCCTCTCCATTATAGGCTTCAACCAACCCAAAGTGTGGCCCAACCCAGAGGTGTTG 1326  
|||||  
DB 1266 TCCTCTCCATTATAGGCTTCAACCAACCCAAAGTGTGGCCCAACCCAGAGGTGTTG 1325  
QY 1327 ACCCTTCCGTTTGTGACCGGGTTGTGTCACACACAGCCACGCTTCTGCTTCTGAG 1386  
|||||  
DB 1326 ACCCTTCCGTTTGTGACCGGGTTGTGTCACACACAGCCACGCTTCTGCTTCTGAG 1385  
QY 1387 GAGATCAAGG 1397  
|||||  
DB 1386 GAGATCAAGG 1396

RESULT 13  
RABCP4A6 1790 bp mRNA linear MAM 27-APR-1993  
LOCUS RABCP4A6 1790 bp mRNA linear MAM 27-APR-1993  
DEFINITION Rabbit cytochrome P4501VA6 (CYP4A6) mRNA, complete cds.  
ACCESSION M28656  
VERSION M28656.1 GI:164976  
KEYWORDS cytochrome P450; lauric acid omega-hydroxylase.  
SOURCE Oryctolagus cuniculus (rabbit)  
ORGANISM Oryctolagus cuniculus

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
1 (bases 1 to 1790)  
Johnson, R. F., Walker, D. L., Griffin, K. J., Clark, J. E., Okita, R. T.,  
Muerhoff, A. S., and Masters, B. S.

TITLE Cloning and expression of three rabbit kidney cDNAs encoding lauric  
acid omega-hydroxylases  
JOURNAL Biochemistry 29 (4), 873-879 (1990)  
MEDLINE 90254128  
PUBMED 2340280

COMMENT Original source text: Rabbit (strain New Zealand White, adult)  
kidney, cDNA to mRNA, clone KdA6.  
Draft entry and computer-readable sequence for [Biochemistry 29,  
873-879 (1989)] kindly submitted  
by E. F. Johnson, 04-OCT-1989.

FEATURES  
Source location/Qualifiers  
1..1790

Source  
/organism="Oryctolagus cuniculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9986"  
14..1546  
/note="lauric acid omega-hydroxylase"  
/codon\_start=1  
/protein\_id="AAA31230.1"  
/db\_xref="GI:164976"

CDs  
/translation="MSVSALNPTRLPGSLGILQVAGLLGLLLKAQULYLRQWL  
LRALQOPCPPEFHWLGHSHREFOHGLQVLMKVEKPSACPRMLWGSRAHLIYDP  
DYMKVILGRSDPKAOGSVRFAPWIGYGLLNGQTFQHRMLTPAFHYDILKPYVG  
LMADSVQIMLDKWEQLVSQDSLEVFQDISLMTLDTIMKCAFSHQGSVQIDRNSQSYI  
QAVGDLNMLFSSRVNRFHQSDTLYRISPGRLSHRAQOLAHEHTDRVIOQRKAQLOQ  
EGELEKVRKRRLDFLDVLLFAKMENGSLSDDLRAEVDTFMFEHDTTASGISWIF  
YALATHPEHQRCREBIQGLLDGASITWEHLDMPTTMCIKBALRYEPVPGVRQ  
LSSPTFPDGRSLPKGVIVTSLIYALHNPKNVNPVEVDFPFPAGSASHSHALP  
SGSPRNCIGKQFAMNBLKVAVALTLVRFBLPDPKRVDPDQKRLVLKSSNGIHLRLRK  
LR"

BASE COUNT 341 a 610 c 479 g 360 t  
ORIGIN

Query Match 48.2%; Score 1241.4; DB 4; Length 1790;  
Best Local Similarity 82.5%; Pred. No. 1.1e-303;  
Matches 1459; Conservative 0; Mismatches 306; Indels 4; Gaps 3;

QY 25 GCTGACCATGAGTGTCTGTGTGTGAGCCCAAGAGACTCCTGGGTGATGTCTTGAA 84  
|||||  
DB 6 GCTGACCATGAGTGTGTGTGTGTGACTGAACCCCAACCGGCTCCGGGACGCTTCCGGGC 65  
QY 85 TCCTCCAGCGGCTCCTGCTCATTTCTGCTTCTGCTGTGATCAAGGCAGTTCAGCTCT 144  
|||||  
DB 66 TCCTCCAGTGGCGGCTGTGCTGGGCTGTCTGCTGTGCTCAAGGCAAGCTCAAGCTCT 125  
QY 145 ACCTGCACAGGCAAGTGGCTGCTCAAGAGCCCTCCAGCAGTTCCTGCTCCCTCCCACT 204  
|||||  
DB 126 ACCTGCACAGGCAAGTGGCTGCTCAAGAGCCCTCCAGCAGTTCCTGCTCCCTCCCACT 185  
QY 205 GGGCTTCCGGGCAATCCAGAGCTCCACAGAGCAAGAGCTACACGGAATTCAGAAAT 264  
|||||  
DB 186 GGGCTTCCGGGCAATCCAGAGCTCCACAGAGCAAGAGCTACACGGAATTCAGAAAT 245  
QY 265 GGGTGAAGACATTCCTCAAGTGTCTGTCTCATTTGGCTATGGGAGGCAAGTTCGTGCC 324  
|||||  
DB 246 GGGTGAAGACATTCCTCAAGTGTCTGTCTCATTTGGCTATGGGAGGCAAGTTCGTGCC 305  
QY 325 AGCTCTATGACCTGACTATATGAAAGTGAATTCGGGAGATCAGACCCGAATCCCATG 384  
|||||  
DB 306 TGATCTATGACCTGACTATATGAAAGTGAATTCGGGAGATCAGACCCGAATCCCATG 365  
QY 385 GTTCTTACAGATTCCTGCTCCTCATGATGATGAGGTACGGCTGTCTCTGTTGAATGGGCAGA 444  
|||||  
DB 366 GTTCTTACAGATTCCTGCTCCTCATGATGATGAGGTACGGCTGTCTCTGTTGAATGGGCAGA 425  
QY 445 CATGTTCCAGCATCCAGCGATGCTGACCCCAAGCTTCCACTATGACATCCTGAAGCCCT 504  
|||||  
DB 426 CATGTTCCAGCATCCAGCGATGCTGACCCCAAGCTTCCACTATGACATCCTGAAGCCCT 485  
QY 505 ATGTGGGCTCATGGCAGACTCTGTACGAGTGAATGCTGACCAATGGGAGAGAGCTCCTTG 564  
|||||  
DB 486 ATGTGGGCTCATGGCAGACTCTGTACGAGTGAATGCTGACCAATGGGAGAGAGCTCCTTG 545  
QY 565 GCCAGATTCCTCTGTGAGGTCTTTACAGACGTCTCTGATGACCTTGAGACACCATCA 624  
|||||  
DB 546 GCCAGATTCCTCTGTGAGGTCTTTACAGACGTCTCTGATGACCTTGAGACACCATCA 605  
QY 625 TGAAGTGTGCTTCAAGCATCAGGCGCAGCATCCAGGTGACAGAAATTCAGTCCCTACA 684  
|||||  
DB 606 TGAAGTGTGCTTCAAGCATCAGGCGCAGCATCCAGGTGACAGAAATTCAGTCCCTACA 665  
QY 685 TACAGGCAATGATGACCTGAACAACCTGTTTTTCCCGTGTGAGAAATGCTTTCAAC 744  
|||||  
DB 666 TACAGGCAATGATGACCTGAACAACCTGTTTTTCCCGTGTGAGAAATGCTTTCAAC 725  
QY 745 AGAATGACACCATCTACAGCCTGACCTGTGCGCGGCTGACACACCGCGCTGCCAGC 804  
|||||  
DB 726 AGAATGACACCATCTACAGCCTGACCTGTGCGCGGCTGACACACCGCGCTGCCAGC 785  
QY 805 TGGCCCATCAGACACAGACCAAGTATCCAATGAGGAAGGCTCAATAAGAGAGAGG 864  
|||||  
DB 786 TGGCCCATCAGACACAGACCAAGTATCCAATGAGGAAGGCTCAATAAGAGAGAGG 845  
QY 865 GGGAGCTGAGAAATCAAGAGGAAGGCAATTTGATTTTCTGAATATCCTCTCTTG 924  
|||||  
DB 846 GGGAGCTGAGAAATCAAGAGGAAGGCTTGAATTTCTGAGCTCTCTCTTG 905  
QY 925 CCAAAATGAGAAATGGAGCATCTGTGACACAGGACCTCCGTGCTGAGGTGACACGT 984  
|||||  
DB 906 CCAAAATGAGAAATGGAGCATCTGTGACACAGGACCTCCGTGCTGAGGTGACACGT 965  
QY 985 TCATGTTGAGGGCCACGACACCAAGGCAAGGATCTCTGATCTCTATGCTCTGG 1044  
|||||  
DB 966 TCATGTTGAGGGCCACGACACCAAGGCAAGGATCTCTGATCTCTATGCTCTGG 1025  
QY 1045 CCACACACCCCAAGCATCAGAGAGGTGCGGGAGAGATCCACGCTCTGGGTATG 1104  
|||||  
DB 1026 CCACACACCCCAAGCATCAGAGAGGTGCGGGAGAGATCCACGCTCTGGGGATG 1085  
QY 1105 GAGCTTCATGACCTGGAACCACTGGAACCAAGATGCGCTTACACCAACATGTGATTAAAG 1164

Db	1086	GAGCCTCCATCACTGGGAGACACCTGGACCAGATGCCCTTACACCAACCATGTGCAATCAAGG	1145
Qy	1165	AGGCACTGAGGCTCTAACCCACCGGTGCCAGGCATTGGCAGAGAGCTCAGCACTCCGCTCA	1224
Db	1146	AGGCGTGAAGACTCTAACCCACGATGCCAGGTGTGGCAGACAGCTCAGCTCACTGTCA	1205
Qy	1225	CCTTCCCTGATGGGCGCTCCTTGCCCAAGGATCATGGTCTCTCTCCATTATGGCC	1284
Db	1206	CCTTCCCTGATGAGCGCTCCCTCCCAAGGATCATAGTCAAGCTCTCCATCTAGGCC	1265
Qy	1285	TTCACCACAACCCCAAAAGTGTGGCCCAACCCAGAGTGTGTGAACCCCTTCCGTTTGAC	1344
Db	1266	TTCACCACAACCCGAAGTGTGGCCAAACCCAGAGTGTGTGAACCCCTTCCGTTTGAC	1325
Qy	1345	CGGTTCTGCTCAACACAGCCACGCTTCTCTGCCCCCTTCTCAGGAGATCAAGAACTGCA	1404
Db	1326	CGGTTCTGCTCGCCACAGCCACGCTTCTCTGCCCCCTTCTCAGGAGACCAAGAACTGCA	1385
Qy	1405	TTGGAAACAATTGGCCATGAAAGAGCTGAAGGTGGCCACGGCCCTGAACCTGTCCGCT	1464
Db	1386	TGGGAAGCAATTGGCCATGAAATGAGCTGAAGGTGGCCCTGAACCTGTCCGCT	1445
Qy	1465	TTGAGCTGCTGCCCTGATCCCAACAGGATCCCATTCGCCATTGTCAGACTTGTGTGAAT	1524
Db	1446	TGAGCTGCTGCCAGATCCCAAAAGAGTCCCGGACCAAAAACCAAGCTTGTGTGCTGAAT	1505
Qy	1525	CCAAAATGGAATCCACTGCGCTCAGAGAGCTCC-CTAACCTTGTGAAGACAAAGAC	1583
Db	1506	CCAGCAACGGGATCCACTGCGCTGAGAGAGCTCCGCTAACCTGTGTGGAGACAAGAC	1565
Qy	1584	CAGCTTTGAGGGCTCCACTGCCGCTCTGTCTTCTGAACCCCGCTTCTGTCCCTTCC	1643
Db	1566	AGGCTCTGGGGCTTCTGTGCCAGGCGTCTGTCTTCTGTCACTGCCCATGTGCCCTGCC	1625
Qy	1644	TGTCTGCCCATATCTGTGTTTCTGTCTGTGCCACCACTTCCCTTCTTCCCACTGCGTGT	1703
Db	1626	TGTCTGCCCATCTGTCTTCTATTCACCAAGCACTTCTTCCACCTGTCTGCTGTGC	1685
Qy	1704	CCCCCAGTGTGACCTGC--CCTTCTCTCTCTACCACTTCTTCCAGGCTCCCTACCTGCTGT	1761
Db	1686	CTCTTGGCCTCCAGGCTGTCTGTCTCTCTCTGCACTTCTCTGTGGGCCACTGACCTGTCTGT	1745
Qy	1762	CTAAGTGTCTCTAACCCCACTGTATCTCT 1790	
Db	1746	CTA-CTGTCCGCTTCTGCAAGCATCTCT 1773	

RESULT 14	
RABCYTR9	
LOCUS	2191 bp
DEFINITION	Rabbit cytochrome P-450-kal mRNA, complete cds.
ACCESSION	M29531 J05150
VERSION	M29531.1 GI:164986
KEYWORDS	Cytochrome P450kal.
SOURCE	Oryctolagus cuniculus (rabbit)
ORGANISM	Oryctolagus cuniculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE	1 (bases 1 to 2191)
AUTHORS	Yokotani, N., Bernhard, R., Sogawa, K., Kusunose, E., Gotoh, O.,
	Kusunose, M. and Fujii-Kuriyama, Y.
TITLE	Two forms of omega-hydroxylase toward prostaglandin A and laurate.
	cDNA cloning and their expression
JOURNAL	J. Biol. Chem. 264 (36), 21665-21669 (1989)
MEDLINE	90094341
PUBMED	2600085
COMMENT	Original source text: Rabbit (Japanese white) adult male kidney, cDNA to mRNA, clone R9.
	Draft entry and printed sequence for [1] kindly submitted by
	N.Yokotani, 27-OCT-1989.
FEATURES	Location/Qualifiers
Source	1..2191

CDS  
 /organism="Oryctolagus cuniculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9986"  
 10.1542  
 /note="cytochrome P-450-kal (EC 1.14.99.)"  
 /codon\_start=1  
 /protein\_id="AAA31234.1"  
 /db\_xref="GI:164987"  
 /translation="MSVSALNPTRLPSLSGLQVAGLLGLLLLLKAAQLYLRHQL  
 LRLAQFPCEPFHMLGHSRBPONHELOVMKVEKPPSACPRMLWGSFAHLITYDE  
 DYMVILGRSDPKAQSRYFLPMWIGYGLLLNGQTWFQHRMLTPAFHYDLKPYG  
 LMADSVQIMLDKWEOLVSQDSLEVFODISLMTLDTIMKCAFSHQSVQDRNSQYI  
 QAVGDIANNLFPSRVNRFHSDTIYRLSPBGRSLSHRACQOLAHETDRIQQRKQLOQ  
 EGBLEKVRNKRRLDFLDVLFAKMNSSLSDDLRAEVDTFMEGHDTYASGISWIE  
 YALATHPEHQHRCREBIQGLLDGASITWEHLDPMPYTTMCKEALRLYPVPGVGRG  
 LSSPTFPDGRSLPKGVITLSIYALHNPKGPNPEVDPSPFAPGSARHSHAFPEP  
 SGGPRNCIGKQFAMNELKVAVALTLRPELLDPKRVDPQKRLVLKSSNGIHLRLRR  
 LR"  
 BASE COUNT 407 a 744 c 558 g 482 t  
 ORIGIN  
 Query Match 47.3%; Score 1219; DB 4; Length 2191;  
 Best Local Similarity 81.7%; Pred. No. 5.5e-298;  
 Matches 1445; Conservative 0; Mismatches 320; Indels 4; Gaps 3;  
 QY 25 GCTGCACCATGATGTCTCTGTGTGCTGAGCCCCCAGAGACTCTGGTGATGTCTTGAA 84  
 DB 2 GCTGCACCATGAGCGTGTCTGCACTGAACCCCACTCCGGGAGCTCTCCGGC 61  
 QY 85 TCCTCCAGCGGCTCCCTGCTCATTTGCTTTGCTGCTGATCAAGCAGTTCACTCT 144  
 DB 62 TCCTCCAGTGGCGGCTCTGCTGGGCTGCTCTGCTGCTGCTCAAGCAGCTCAGCTCT 121  
 QY 145 ACCTGCACAGGCACTGGCTGCTCAAGCCCCCTCCAGCAGTTCCCGTGCCCTCCCACT 204  
 DB 122 ACCTGCACCGCCAGTGGCTGCTCAGAGCCCTCCAGCAGTTCCCGTGCCCACTCCACT 181  
 QY 205 GGCTCTTCGGGCACATCCAGAGAGCTCCACACAGACAGAGCTACAAGGATTGAGAAAT 264  
 DB 182 GGCTCTTCGGGCACAGCGAGAGATTCCAAATAGGCCATGAGTTAACAAGTATGTAAT 241  
 QY 265 GGGTGAAGACATTCCTCAAGTGCCTGCTCATTTGGCTATGGGAGAGCAAAAGTTGCTGCC 324  
 DB 242 GGGTGAAGAAATTCCTCAAGTGCCTGCTCATTTGGCTATGGGAGAGCAAGCCACTCC 301  
 QY 325 AGCTCTATGACCTGACTATATGAAGGTGATTCTGGGGAGATCAGACCCGAATCCCATG 384  
 DB 302 TGATCTATGACCTGACTATATGAAGGTGATTCTGGGGAGATCAGACCCGAAGCTCAAG 361  
 QY 385 GTTCTCTACAGATTCCTGCTCCATGATGGGTACCGGCTTGCTCTGTTGAATGGGCAGA 444  
 DB 362 GTTCTCTACAGATTCCTGCTCCCTGATGGGTATGGTTGCTCTGCTGATGGGCAGA 421  
 QY 445 CATGTTCCAGCATGCAAGGATGCTGACCCCAAGCTTCCACTATGACATCTCTGAAGCCCT 504  
 DB 422 CGTGTTCAGCACCGGCGCATGCTCAACCCAGCTTCCACTACGACATCTCTGAAGCCCT 481  
 QY 505 ATGTGGGGCTCATGGCAGACTCTGTACGAGTGTGCTGACAAATGGGAAGAGCTCTTG 564  
 DB 482 ACGTGGGGCTCATGGCGGACTCTCGTCCAAATCATGCTGACAAATGGGAAGAGCTGTCA 541  
 QY 565 GCCAGGATTCCTCTGAGAGTCTTTCAGCAGCTCTCTTGAATGACCCCTGACACCATCA 624  
 DB 542 GCCAGGACTCTCTCTGAGAGTCTTTCAGCAGCATCTCTGATGACCCCTGACACCATCA 601  
 QY 625 TGAAGTGTGCTTCAAGCATCAGGGCAGCATCCAGGTGACAGGAATCTCAGTCTTACA 684  
 DB 602 TGAAGTGTGCTTCAAGCATCAGGGCAGCGTCCAGTTGACAGGAATCTCAGTCTTACA 661  
 QY 685 TACAGGCCATTAGTGACTTGAACAACCTGTTTTTCCCGTGAAGGAATGCTTTCAAC 744  
 DB 662 TCCAGGCTGTTGGGAGCTTGAACAACCTGTTTTTCCCGGAGTGAAGGAATGCTTTCAAC 721

QY	745	AGATGACACCATCTACAGCCTGACCTCTGTCGGCCGCTGACACACCGCGCCTGCAGC	804
Db	722	AGAGTGACACCATCTACAGCCTGAGCCCTGAAAGCCGCTTGTCACCGTGCCTGCAGC	781
QY	805	TGCCCCATCAGCACACAGACCAAGTGATCCAACTGAGGAAGCTCACTACAGAAAGAGG	864
Db	782	TGCCCCAGACACACAGACCGAGTGATCCAGCAGAGGAAGGCTCAGCTGCAGCAGAGG	841
QY	865	GGAGCTGGAAGATCAAGAGGAAGGCAATTTGGATTTTCTGATATCCTCCTTGG	924
Db	842	GGGAGCTGGAAGAGGTCAAGAGGAAGAGGCGCTTGGACCTTCTGACGTCCTCTTGG	901
QY	925	CCAAATGGAATGGAGCACTTGTCAACAAGGACCTCCGTGCTGAGGTGACACGT	984
Db	902	CCAAATGGAATGGAGCAAGCTGTCCGACAGGACCTCCGCGCGAGGTGACACGT	961
QY	985	TCATGTTGAGGGCCACGACACACAGCCAGTGGGATCTCTGAGTCTCTATGCTTGG	1044
Db	962	TCATGTTGAGGGCCACGACACACAGCCAGCGGATCTCTGAGTCTCTATGCTTGG	1021
QY	1045	CCACACACCCCAAGCATCAGAGAGGTGCCGGAGAGATCCACAGCCTCCTGGTGATG	1104
Db	1022	CCACGACCCCGAGCATCAGCACCGGTGCCGGAGAGATCCAGGGCCTCCTGGGGAGG	1081
QY	1105	GAGCCTCCATCACTTGGAAACCACTGGAACCAAGATGCCCTTACACCACTATGCAATTAAG	1164
Db	1082	GAGCCTCCATCACTTGGAGCACTTGGACCAAGATGCCCTTACACCACTATGCAATTAAG	1141
QY	1165	AGGCACTGAGGCTCTACCCACCGGTGCCAGGCATTGGCAGAGACTCAGCACTCCCGTCA	1224
Db	1142	AGGCGTGAGACTTACCCACAGTGCCAGGTGTGGCAGACAGCTCAGCTCACCTGTCA	1201
QY	1225	CCTTCCCTGATGGGCGCTCCTTGCCCAAGGATCATGCTCTCTCCATTATGAGCC	1284
Db	1202	CCTTCCCTGATGAGCGCTCCCTCCCAAGGATCATGCTCTCCATTATGAGCC	1261
QY	1285	TTCAACCAACCCAAAAGTGTGGCCCAACCCAGAGGTGTTGAACCTTCCGTTTGGAC	1344
Db	1262	TTCAACCAACCCGAAGTGTGGCCAAACCCAGAGGTGTTGAACCTTCCCGTTGGAC	1321
QY	1345	CGGGTCTGCTCAACAACAGCCAAGCTTTCCTGCGCTTCTCAGAGGATCAAGGAATGCA	1404
Db	1322	CGGGTCTGCTCGCCACAGCCAAGCTTTCCTGCGCTTCTCAGAGGATCAAGGAATGCA	1381
QY	1405	TTGGAAAACAATTTGCCATGAAAGAGCTGAAGGTGGCCACGGCCCTGACCCTGCTCGCT	1464
Db	1382	TGCGGAAGCAATTTGCCATGAAAGAGCTGAAGGTGGCGGTGACCTGCTGCGCT	1441
QY	1465	TTGAGCTGCTGCTGATCCCAACAGGATCCCCCATCCCCATGCAAGACTTGTGTAAT	1524
Db	1442	TGAGCTGCTGCTGATCCCAAAAGATCCCGGACCAAAACCAAGCTTGTGCTGAAT	1501
QY	1525	CCAAAAATGGAATCCAAGCTGCGTCTCAGGAGGCTCC-CTAACCTTGTGAAGACAAGAC	1583
Db	1502	CCAGCAACGGGATCCAAGCTGCGTCTGAGGAAGCTCCGCTAACCTTGTGGGGAACAAGAC	1561
QY	1584	CAGCTTGAGGGCTCCACAGCTGCGCTCTGCTTCTTGACCCCGCTTCTGTCCCTTCC	1643
Db	1562	AGGCTTGAGGGCTTCTGCAAGGCGTCTCGTCTCTGCACTGCAAGCCCTGCGG	1621
QY	1644	TGTCGCCCATATCCTGTTTCTGTCTGCCACCTTCCCTTCTTCCACCTGCGCTGT	1703
Db	1622	TGTCGCCCATATCCTGCTTCTATCCACAGCACTTCTTCTCTGTCTGCTTGTCTGC	1681
QY	1704	CCCCCACTGCTGCCTG--CCTTCTCTCTCACTTCTTCCAGGCTCCCTACCTGCTGT	1761
Db	1682	CTCTTGGCCTCCAGGCTGTCTGTCTCTCTGCAACCTTCTTGGGCACTGACCTGTCTGT	1741
QY	1762	CTACCTGTCTCTACCCCACTGTATCTCT 1790	
Db	1742	CTA-CTGTCCGCTTCTGCAAGATCTCT 1769	

RESULT 15	OCPP450KD	2377 bp	mrna	linear	MAM 22-MAY-1991
LOCUS	OCPP450KD				
DEFINITION	Rabbit mRNA for omega-hydroxylase cytochrome P-450(kd).				
ACCESSION	X57209				
VERSION	X57209.1	GI:1655			
KEYWORDS	Cytochrome P450.				
SOURCE	Oryctolagus cuniculus (rabbit)				
ORGANISM	Oryctolagus cuniculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Lagomorpha; Leporidae; Oryctolagus.				
AUTHORS	1 (bases 1 to 2377)				
TITLE	Yokotani,N., Kusunose,E., Sogawa,K., Kawashima,H., Kinoshaki,M., Kusunose,M. and Fujii-Kuriyama,Y.				
JOURNAL	CDNA cloning and expression of the mRNA for cytochrome P-450kd which shows a fatty acid omega-hydroxylating activity				
MEDLINE	Eur. J. Biochem. 196 (3), 531-536 (1991)				
PUBMED	91192021				
REFERENCE	2013275				
AUTHORS	2 (bases 1 to 2377)				
TITLE	Fujii-Kuriyama,Y.				
JOURNAL	Direct Submission				
FEATURES	Submitted (28-DEC-1990) Y. Fujii-Kuriyama, DEPT OF BIOCHEMISTRY, FACULTY OF SCIENCE, TOHOKU UNIVERSITY, ARAMAKI AOBIA-KU, SONDAI 980, JAPAN				
source	Location/Qualifiers				
gene	1. .2377				
mRNA	/organism="Oryctolagus cuniculus"				
CDS	/mol_type="mRNA"				
	/db_xref="taxon:9986"				
	/cissue_type="kidney"				
	1. .2377				
	/gene="P-450 (kd) "				
	1. .2377				
	/gene="P-450 (kd) "				
	/evidence=experimental				
	22. .1557				
	/gene="P-450 (kd) "				
	/codon_start=1				
	/product="omega-hydroxylase cytochrome P-450"				
	/protein_id="CAA40493.1"				
	/db_xref="GI:1656"				
	/db_xref="SWISS-PROT:P14579"				
	/translation="MSVSALSPTRLPGSLGLQVAAALLGILLLLKAAQLYLRRQWL LRALQFPCCPFHMLLGHSHREFQNGELQIILKWEKFRACPHWIGNKVRVQLYDP DVMKVTILGRSDPKSRGTYTFVAPWIGVGLLLNGQPFQHRMLTPAFHYDILKPYVG LAMVSVQIMLDKWEQVLSQDSLEVFQDISLMTLDTIMKCAFSQGSVOLDSRNSQSY IQAVGDNLNLVFARVRNIFHOSPTIYRLSPEGRLSHRQAQHAHTDRVIOQRKAQLO QEGELKVRKRRLDFLDVLFLPAKENGSSLSDDDLRAEVDPTFMEGHDTASGVSWI FYALATPEHQHRCREBIQGLGDGASITWEHLQMPYTTWCIKEARLYPPVPAISR DLSSEVTEPDDGSLPKGFTVTLSIYGLHNPVWPNPEVFPDSRFTPGSARHSHAFLP FSGGARNCIGKQFAMNBLKVAVALTLRLFELLPPTRIIPKPTARLVLSNNGIHRLRL KLO"				
	2342. .2347				
	/gene="P-450 (kd) "				
	758 c				
	565 g				
	557 t				
BASE COUNT	497 a				
ORIGIN					
polyA_signal					
Query Match	46.5%; Score 1197.2; DB 4; Length 2377;				
Best Local Similarity	83.4%; Pred. No. 1.9e-292;				
Matches 1396; Conservative	0; Mismatches 273; Indels 5; Gaps 13;				
DB	21 AGGTGTCACCATGAGTGTCTCTGTGTGAGCCCCAGCAGACTCTGGGTGATGTCTCT 80				
	10 AGCGGTGACCATGAGCGTGTCTGACATGAGCCCAACCGGCTCCCGGAGCCTCTCC 69				
OY	81 GGAATCTTCAAGCGGCTCCCTGTCTCATTTCTGTTTGTGTGATCAAGCAGTTTCA 140				
DB	70 GGGCTCTTCAAGTGGCGGCTGTGGGCTCTCTGTCTGTCTCAAGCAGCTCAG 129				
OY	141 CTCTACGTGACAGGCAAGTGGCTGCTCAAGCCCTCCAGCAGTTCCCGTCCCTCC 200				
DB	130 CTCTACGTGCGCCGACGATGGCTGTCTCAGAGCGCTCCAGCAGTTCCCGTCCCGCCCTTC 189				



QY 201 CACTGCTCTTGGGACATCCAGAGCTCCAAAGAGACAGAGCTCAACGGATTGAG 260  
DB 190 CACTGCTCTTGGGACAGCCGAGAGTCCAGATGAACAGAGAGCTCAACAGATTCTG 249  
QY 261 AAATGGGTGAGACATTTCCCAAGTGCCTGCTCATTTGCTATGGGAGGCAAGTTGCT 320  
DB 250 AAATGGGTGAGAGAGTCCCAAGAGCTGCTCATTTGATAGGGGCAACAAGTCCGA 309  
QY 321 GTCCAGCTCTATGACCTCTGACTATATGAAGTGAATTCTGGGAGATCAGACCCGAATCC 380  
DB 310 GTCCAGCTCTATGACCCGGACTACATGAAGTGAATCTGGGAGATCAGATCCAAAGTCT 369  
QY 381 CATGCTCTTACAGATTCTGCTCCATGATTTGGGTACGGCTTGTCTGTTGAATGG 440  
DB 370 CGTGGTCTTACATCTGCTGCTCCCTGAGTTGGGTACGGTGTGCTCTGCTGAACGG 429  
QY 441 CAGACATGTTCCAGCATCCAGAGTGAACCCAGCCTTCCATATGACATCTCTGAAG 500  
DB 430 CAGCATGTTCCAGACCCGCGCATGCTCAACCCAGCCTTCCATACGACATCTCTGAAG 489  
QY 501 CCCTATGCTGGGCTCATGCGCAGACTCTGTAAGATGATGCTGAGCAATGGGAAGCTC 560  
DB 490 CCCTATGCTGGGCTCATGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 549  
QY 561 CTTGGCCAGAGTCCCTCTGAGGCTTTTCAAGCAGCTCTCTGATGACCTTGAACACC 620  
DB 550 GTGAGCCAGAGTCTCTCTGAGGCTTCTCAAGACATCTCTCTGATGACCTTGAACACC 609  
QY 621 ATCATGAAGTGTGCTTACGACCATCAGGCGCAGATCCAGGTGA---CAGGAATTTCTAG 677  
DB 610 ATCATGAAGTGTGCTTACGACCATCAGGCGCAGCTGAGTGAACAGGAATTTCCAG 669  
QY 678 TCCTACATACAGGCGCATTAAGTACCTGAACAACTGTTTTCCTGCTGAGGAATGCC 737  
DB 670 TCCTACATACAGGCGCTTGGGAGCTGAACAACTGTTTTCCTGCTGAGGAATGCC 729  
QY 738 TTTCAACGAGATGACCATCTACAGCTGACTCTGCTGCGCTGAGACACACCGCGCC 797  
DB 730 TTTCAACGAGCGACACCATCTACAGGCTGAGCTTGAAGGCGCTTGTCCACCGTGC 789  
QY 798 TGCAGCTGGCCATCAGCACACAGACCAAGTATCCAACTGAGGAAGGCTCAACTACAG 857  
DB 790 TGCAGCTGGCCATCAGCACACAGACCAAGTATCCAACTGAGGAAGGCTCAGTGCAG 849  
QY 858 AAGAGGGGAGCTGAGAAAGATCAAGAGGAGAGGCAATTGATTTTCTGATATCTCTC 917  
DB 850 CAGAGGGGAGCTGAGAAAGTCAAGAGGAGAGGCGCTTGAATCTCTGAGCTCTCTC 909  
QY 918 CTCTGGCCAAAATGAGAGATGGAGCATCTGTCAACAAGACCTCCGCTGAGGTG 977  
DB 910 CTCTGGCCAAAATGAGAGATGGAGCATCTGTCAACAAGACCTCCGCGCGAGGTG 969  
QY 978 GACACGTTCAATGTTGAGGGGCCACAGACCAAGCAGTGGATCTCTGATCTCTAT 1037  
DB 970 GACACGTTCAATGTTGAGGGGCCACAGACCAAGCAGTGGATCTCTGATCTCTAT 1029  
QY 1038 GCTCTGGCCACACCCCAAGCATCAGAGAGGTGCGGGAGAGATCAGAGCTCTCTG 1097  
DB 1030 GCTCTGGCCACACCCCAAGCATCAGAGAGGTGCGGGAGAGATCAGAGCTCTCTG 1089  
QY 1098 GGTGATGAGCTCTCACTACCTGGAACCACTGGAACAGATGCCCTACCACTGATGC 1157  
DB 1090 GGTGATGAGCTCTCACTACCTGGAACCACTGGAACAGATGCCCTACCACTGATGC 1149  
QY 1158 ATTAAGAGGCACTGAGCTTACCCACCGGTGAGGATTTGAGAGAGCTCAGCACT 1217  
DB 1150 ATTAAGAGGCGATGAGCTTACCCACCGGTGAGGATTTGAGAGAGCTCAGCACT 1209  
QY 1218 CCGCTCACTTCCCTGATGGGCGCTCTTGGCCAAAGTATCATGCTCTCTCTCATTT 1277  
DB 1210 CCGCTCACTTCCCTGATGGGCGCTCTTGGCCAAAGTATCATGCTCTCTCTCATTT 1269

QY 1278 TATGCTTCAACCAACCCAAAAGTGTGCCCCAACCCAGAGGTGTTGACCTTTCCGT 1337  
DB 1270 TATGCTTCAACCAACCCGAATGTGTGCCCCAACCCAGAGGTGTTGACCTTTCCGC 1329  
QY 1338 TTTGACCGGTTCTGCTCAACACAGCCAGCTTTCTGCTGCTTTCTCAGAGATCAAG 1397  
DB 1330 TTTGACCGGTTCTGCTCAACACAGCCAGCTTTCTGCTGCTTTCTCAGAGGAGCAAG 1389  
QY 1398 AACTGATTTGGGAAACAATTTGCCATGAAGAGCTGAAGTGGCCACGCTTGAACCTG 1457  
DB 1390 AACTGATTTGGGAAACAATTTGCCATGAAGAGCTGAAGTGGCCACGCTTGAACCTG 1449  
QY 1458 CTCCGCTTGAAGTGTGCTGCTGATCCACAGATCCCATTTCCATTTGACGACTGTG 1517  
DB 1450 CTCCGCTTGAAGTGTGCTGCTGATCCACAGATCCCATTTCCATTTGACGACTGTG 1509  
QY 1518 TTGAATCCAAAATGGAATCCACCTGCTCTCAGAGGCTCC-CTAACCTTGTGAAGA 1576  
DB 1510 TTGAATCCAAAATGGAATCCACCTGCTCTCAGAGGCTCC-CTAACCTTGTGAAGA 1569  
QY 1577 CAAGACCAAGCTTGAAGGCTCC-ACCTGCGCTCTGTCTTCTGACCCCGCTTCTGT 1635  
DB 1570 CAAGACCAAGCTTGAAGGCTCC-ACCTGCGCTCTGTCTTCTGACCCCGCTTCTGT 1629  
QY 1636 CCGCTTCTGCTGCTGCTATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1689  
DB 1630 CCGCTTCTGCTGCTGCTATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1683

Search completed: February 15, 2004, 03:56:54  
Job time : 9395 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 14, 2004, 16:07:28 (Search time 672 Seconds  
(without alignments)  
10347.846 Million cell updates/sec

Title: US-09-942-429A-6

Perfect score: 2576  
Sequence: 1 gaattccgcagagatccagc.....aaaaaaaaaacggaattc 2576

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq.19Jun03:\*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2576	100.0	2576	24	ABN95640 Gene #2138 used to
2	2576	100.0	2576	24	ABK50069 DNA encoding human
3	2576	100.0	2576	24	ABL68591 Kidney cancer rela
4	1625.8	63.1	1763	24	ABN95112 Gene #1610 used to
5	1625.8	63.1	1763	24	ABL68271 Kidney cancer rela
6	1625.8	63.1	1763	24	ABL68599 Kidney cancer rela
7	1625.8	63.1	1763	24	ABL68881 Kidney cancer rela
8	1525.2	59.2	1872	24	ABK50070 DNA encoding human

9	1082.8	42.0	2116	24	ABK50068	DNA encoding mouse
10	1022.4	39.7	2462	24	ABK63715	Rat sequence diffe
11	968.4	37.6	21990	24	ABK50071	Human Cyp 4A11 gen
12	966	37.5	4123	24	ABK50067	DNA encoding mouse
13	534	20.7	2343	22	AAS21297	Human CDNA sequenc
14	534	20.7	2343	25	ACA03656	CDNA encoding huma
15	534	20.7	2343	25	ACA04077	Human CDNA encodin
16	534	20.7	2343	25	ABX89194	DNA encoding novel
17	534	20.7	2381	24	ABK63921	CDNA encoding huma
18	534	20.7	2416	22	AAD09941	Human drug metabol
19	528.8	20.5	1894	24	ABT07693	Breast cancer-asso
20	528.8	20.5	2015	24	AAD24015	Human drug metabol
21	528.8	20.5	2020	24	ABK33550	CDNA encoding huma
22	519.4	20.2	2327	24	ABK52898	CDNA encoding huma
23	512.4	19.9	2356	22	AB199563	Human expressed po
24	512.4	19.9	2356	22	ABA06504	Human CDNA SEQ ID
25	512.4	19.9	2356	22	AAS40801	CDNA encoding nove
26	512.4	19.9	2356	22	AAS28959	CDNA encoding for
27	512.4	19.9	2356	22	AAS29592	Human endocrine po
28	512.4	19.9	2356	22	AAS34859	CDNA encoding nove
29	512.4	19.9	2356	24	ABV83841	Human polynucleoti
30	512.4	19.9	2356	24	ABT07794	Novel human nuclei
31	512	19.9	1718	24	ABZ11538	Human polynucleoti
32	506.8	19.7	2084	24	ABL67800	Oesophagus cancer
33	408.4	15.9	1586	22	AAD09946	Human drug metabol
34	392.2	15.2	1657	25	ACC50320	Breast cancer asso
35	387.4	15.0	17717	22	AAK82007	Human immune/haema
36	387.4	15.0	17717	22	AAK82008	Human immune/haema
37	387.4	15.0	17719	22	AAK82006	Human immune/haema
38	359.8	14.0	1508	24	ABZ11227	Human polynucleoti
39	320	12.4	433	25	ABX45374	Bovine EST associa
40	318.4	12.4	2073	24	ABK83594	Human CDNA differe
41	318.4	12.4	2073	24	ABN95103	Gene #1601 used to
42	318.4	12.4	2073	24	ABL68352	Kidney cancer rela
43	318.4	12.4	2073	24	ABL68820	Kidney cancer rela
44	315.2	12.2	2211	23	AAS67737	DNA encoding novel
45	315.2	12.2	2211	23	AAS93586	DNA encoding novel

ALIGNMENTS

ABN95640	ABN95640 standard; DNA; 2576 BP.	
XX	ABN95640;	
AC	13-AUG-2002 (first entry)	
XX		
DT	Gene #2138 used to diagnose liver cancer.	
XX		
DE	Gene, liver cancer; dr; hepatocellular carcinoma; hepatotropic;	
XX	KW metastatic liver tumour; cytostatic; expression profile; disease state;	
KW	disease progression; drug toxicity; drug efficacy; drug metabolism.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200229103-A2.	
XX		
PD	11-APR-2002.	
XX		
PF	02-OCT-2001; 2001WO-US30589.	
XX		
PR	02-OCT-2000; 2000US-237054P.	
XX		
PA	(GENE-) GENE LOGIC INC.	
XX		
PI	Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;	
XX	WPI; 2002-426119/45.	
DR		
XX	Diagnosing and detecting the progression of liver cancer,	
PT		

PT hepatocellular carcinoma or metastatic liver tumor in a patient,  
PT involves detecting the level of expression of two or more genes in a  
PT liver tissue sample -

PS Claim 1; SEQ ID NO 2138; 298bp; English.

The invention relates to a novel method for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumour in a patient, and differentiating metastatic liver cancer from hepatocellular carcinoma in a patient, involving detecting the level of expression of two or more genes represented in ABN93503-ABN97455 in a tissue sample. The method of the invention has hepatotropic, and cytosolic activity. The method is useful for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma and metastatic liver carcinoma in a patient. The method is useful for identifying expression profiles which serve as useful diagnostic markers as well as markers that can be used to monitor disease states, disease progression, drug toxicity, drug efficacy and drug metabolism.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published/pct/sequences](http://ftp.wipo.int/pub/published/pct/sequences).

Sequence 2576 BP; 572 A; 793 C; 583 G; 628 T; 0 other;

Query Match	100.0%;	Score 2576;	DB 24;	Length 2576;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2576;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 GAATTCGAGAGATCCAGCAGGTGCTGCACCATGAGTCTCTGTGCTGAGCCCGACGA 60

Db 1 GAATTCGAGAGATCCAGCAGGTGCTGCACCATGAGTCTCTGTGCTGAGCCCGACGA 60

Oy 61 GACTCTGGGTGATGTCTCTGGAATCTCCAGGGCTCCCTGCTCATTTCTGTTCTGC 120  
 |||||  
 Db 61 GACTCTGGGTGATGTCTCTGGAATCTCCAGGGCTCCCTGCTCATTTCTGTTCTGC 120

121 TCGTATCAAGCAGTTCAGCTCTACCTGACAGGAGTGGCTGCTCAAAGCCCTCCAGC 180  
 121 TCGTATCAAGCAGTTCAGCTCTACCTGACAGGAGTGGCTGCTCAAAGCCCTCCAGC 180  
 121 TCGTATCAAGCAGTTCAGCTCTACCTGACAGGAGTGGCTGCTCAAAGCCCTCCAGC 180

181 AGTTCCTGCTCCCTCCCACTGGCTTTCGGGCATCCAGAGCTCCACAGACC 240

241 AGAGCTACAAACGATTTCAGAAATGGGTGAGACATTTCCCAAGTGCCTGTCTCATTTGGC 300

301 TATGGGAGGCAAGTTGTCGTCCAGCTCTATGACCTGACTATATGAGGTGATTCGG 360

361 GGAGATCAGACCCGAATCCCATGTTCTCTACAGATTCTGGCTCCATGATGGGTACG 420

421 GCTTGCTCCTGTTGAATGGGCAAGACATGGTTCAGCATCGACGGATGCTGACCCCAAGCT 480

481 TCCACTATGACATCCTGAAGCCCTATGTGGGGCTCATGGCAGACTCTGTACGAGTGAATGC 540

541 TGGACAAATGGGAAGAGCTTCCTTGGCCAGGATTCCCCCTCTGAGGGTCTTACGACAGTCT 600

601 CCTTGATGACCCCTGGACACCATCATGAGTGTGCCTCAGCCATCAGGGCAGCATCCAGG 660

661 TGGACAGGAATTCAGTCTACATACAGGCCATTAGTGAACCTGAACAACCTGTTTTT 720

Db 661 TGGACAGGAATTCAGTCCCTACATACAGCCATTAGTACCTGAACAACCTGGTTTTT 720

```

Oy      721 CCCGTGTGAGGAATGCCCTTCA CCAAGATGACACCATCTACAGCCTGACCTCTGCTGGCC 780
      |||||
Db      721 CCCGTGTGAGGAATGCCCTTCA CCAAGATGACACCATCTACAGCCTGACCTCTGCTGGCC 780

```

QY	781	GCTGGACACACCGCGCCCTGCAGCTGGCCCATCAGCACACAGACCAAGTATCCAACTGA	840
DB	781	GCTGGACACACCGCGCCCTGCAGCTGGCCCATCAGCACACAGACCAAGTATCCAACTGA	840

```

Ox      841 CGAAGGCTCAACTACAGAAGCGGGGAGCTGGAGAAGATCAAGAAGAGGCAATTTGG 900
      |||||
Db      841 GGAAGGCTCAACTACAGAAGCGGGGAGCTGGAGAAGATCAAGAAGAGGCAATTTGG 900

```

Qy		960
	901 ATTTTCGATATCCTCTCTTGGCCAATAAGAGAATGGAGCATCTTGTCAACAAG	
Db	901 ATTTTCGATATCCTCTCTTGGCCAATAAGAGAATGGAGCATCTTGTCAACAAG	960

QY 961 ACCTCCGTGCTGAGGTGACACGTTTCATGTTTGAGGGCCACGACACCAAGCCAGTGGGA 102

Db 961 ACCTCCGTGCTGAGGTGACACGTTTCATGTTTGAGGGCCACGACACCAAGCCAGTGGGA 102

Dy	1021 TCTCCTGGATCCTCTATGCTCTGGCCACACCCCAAGCATCAGAAGGTGCCGGAGG	108
Db	1021 TCTCCTGGATCCTCTATGCTCTGGCCACACCCCAAGCATCAGAAGGTGCCGGAGG	108

QY	1081	AGATCCACAGCCTCCTGGTGATGAGCCTCATCACCTGGAACCACTGACCAATGC	114
Dh	1081	AGATCCACAGCCTCCTGGTGATGAGCCTCATCACCTGGAACCACTGACCAATGC	114

QY	1141	CTTACACCA	CCATGTGCAT	TAAAGAGCA	CTGAGGCT	TACCCACCG	GTGCAGCA	TTG	120
nb	1141	CTTACACCA	CCATGTGCAT	TAAAGAGCA	CTGAGGCT	TACCCACCG	GTGCAGCA	TTG	120

QY 1201 GCAGAGAGCTCAGCACTCCCGTCACTTCCCTGATGGGGCTCCTTGCCCAAAGTATCA 1206  
 |||||  
 Db 1201 GCGAGAGGCTCAGCACTCCCGTCACTTCCCTGATGGGGCTCCTTGCCCAAAGTATCA 1206

QY		1261	TGCTCCTCTCCATTATGGCGCTTACACAACCAGAAAGTGCGCCCAACCCAGAGG	1322
B		1262	TTCTTCCTCCTCTCCATAATTAATGCCTTCAACAACCCAAAAGTAGTGCGCCCAACCCAGAGG	1323

1321 TGTTGACCCCTTCGGTTTGACCGGGTCTGCTCAACACAGCCAGCTTTCCTGCCCC 136

1381 TCTCAGGAGATCAAGGACTGCATTGGAAACAATTGCCATGACGAGCTGAAGGTGG 1444

1441 CCACGACCTGACCCCTGCTCCGCTTGAGCTGTGCTGATCCACAGATCCCCATCC 1560

1501 CCATTGACGACTTGTGTGAATCCAAAATGGAATCCACCTGGCTCAGAGGCTCC 1502

1561 CTACCTTTGAGACAAGACCAGCTTTGAGGCGCTCCACCTGCGTCTGTCTTCT 16

1621 GACCCGCGCTCTGTGCCCCCTTCCCTGTCTGCCCATATCTGTCTGTCTGCCCACTTC 1622

1681 CCTTCCTCCACCTGCTGCTGTCGCCCAAGTCGCGCTGCCCTTCTCTCTCTCACCTTTCT 17

1741 CCAAGTCCCTACCTGCTGTCTACCTGTCTCTACCACTGTATCTTGTGGAGA 18

Db 1741 CCAGGCTCCCTACTGCTTGTCTACTGTCTCTAACCCAGCTGTATCTTGTGGAGA 1800  
QY 1801 AAAGCTGAGTGTGGAGAGAGCTGAGGCCGACCTTGACATGTCTGAATAATGTAAGAG 1860  
Db 1801 AAAGCTGAGTGTGGAGAGAGCTGAGGCCGACCTTGACATGTCTGAATAATGTAAGAG 1860  
QY 1861 TCTTGAATCATGTCCAGGATCCAGGGCTTAACCCCTTGTGGCTTTGGAACACCAAGC 1920  
Db 1861 TCTTGAATCATGTCCAGGATCCAGGGCTTAACCCCTTGTGGCTTTGGAACACCAAGC 1920  
QY 1921 TCTGTGCTGAAGGCTGGAAGGCTACCTGACGACCATTAATCTAAGCCCGGACATAAAA 1980  
Db 1921 TCTGTGCTGAAGGCTGGAAGGCTACCTGACGACCATTAATCTAAGCCCGGACATAAAA 1980  
QY 1981 CCCCTGCTGGCTTGATAGATCCAGGGCTCGTGGCTCTGGAATGTCTGACTTGGCTG 2040  
Db 1981 CCCCTGCTGGCTTGATAGATCCAGGGCTCGTGGCTCTGGAATGTCTGACTTGGCTG 2040  
QY 2041 GCTCCTTGCTCTTGTCTCTCCAGGATCAATTTATCTTGAGTTAAAGAACCTGCTCTC 2100  
Db 2041 GCTCCTTGCTCTTGTCTCTCCAGGATCAATTTATCTTGAGTTAAAGAACCTGCTCTC 2100  
QY 2101 CATTATCTCAAGTAAAGAGCAGAGCAGATGTAAACCGTACAGCTGTAAATTTGTGCTTAAT 2160  
Db 2101 CATTATCTCAAGTAAAGAGCAGAGCAGATGTAAACCGTACAGCTGTAAATTTGTGCTTAAT 2160  
QY 2161 GCAACATGCCCTTTCGACCCACCCCAATCTACACCACTGTTCTTGTGTTGATCAGCA 2220  
Db 2161 GCAACATGCCCTTTCGACCCACCCCAATCTACACCACTGTTCTTGTGTTGATCAGCA 2220  
QY 2221 ATAATAATCTGCACTTCCAGAGCTCGGGGCTTACAGCCTCCATCTTAGCTTTGGCG 2280  
Db 2221 ATAATAATCTGCACTTCCAGAGCTCGGGGCTTACAGCCTCCATCTTAGCTTTGGCG 2280  
QY 2281 CCCTGAGCCCACTTCTCTCTCAAACTGTCTTTCTCACTGCTTTGAATCTGCGGACTT 2340  
Db 2281 CCCTGAGCCCACTTCTCTCTCAAACTGTCTTTCTCACTGCTTTGAATCTGCGGACTT 2340  
QY 2341 TGTCAACCCCAAGCAGCTGTGTGGTCTGAACACCCCAATCTCTGAATCTCCACCCA 2400  
Db 2341 TGTCAACCCCAAGCAGCTGTGTGGTCTGAACACCCCAATCTCTGAATCTCCACCCA 2400  
QY 2401 CCTCCCAAACTCTGCTGCTGCCCTCCAGACTGTCTGCCATACACCTGTCTCTCTCT 2460  
Db 2401 CCTCCCAAACTCTGCTGCTGCCCTCCAGACTGTCTGCCATACACCTGTCTCTCTCT 2460  
QY 2461 GCCTGGCTGTCTGTCTCTATATTAGTTCTCTATTACTGCTGTAAATACTATCACAATC 2520  
Db 2461 GCCTGGCTGTCTGTCTCTATATTAGTTCTCTATTACTGCTGTAAATACTATCACAATC 2520  
QY 2521 TCAGTGAATTTAAATAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 2576  
Db 2521 TCAGTGAATTTAAATAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 2576

RESULT 2  
ABK50069  
ID ABK50069 standard; DNA; 2576 BP.

XX ABK50069;  
DT 15-JUL-2002 (first entry)

DE DNA encoding human Cyp 4A11 protein.

KM Human; Cyp 4A14; hypertension; Cyp 4A11; testosterone inhibitor; obesity;  
lipid metabolism disease; pancreatic dysfunction; type II diabetes;  
cardiovascular disease; gene; ds.

OS Chimeric - Homo sapiens.  
OS Chimeric - Synthetic.

XX Key Location/Qualifiers

FT CDS 33.1592  
FT /\*tag= a  
FT /product= "Human Cyp 4A11 protein"  
XX WO200217856-A2.  
XX 07-MAR-2002.  
XX 29-AUG-2001; 2001WO-US26914.  
XX 29-AUG-2000; 2000US-228947P.  
XX (UYVA-) UNIV VANDERBILT.  
XX Capdevila J, Waterman M, Holla V;  
XX WPI; 2002-382929/41.  
XX P-PSDB; AAU79994.  
PT Treating hypertension in an individual by inhibiting testosterone  
PT activity, enhancing Cyp 4A14 activity, or by inhibiting Cyp 4A11  
PT activity in the individual -  
XX Claim 11; Page 71-72; 80pp; English.  
XX The present invention relates to a new method of treating hypertension in  
XX an individual. The method comprises inhibiting testosterone activity,  
XX enhancing Cyp 4A14 activity, or inhibiting Cyp 4A11 activity in the  
XX individual. Enhancing Cyp 4A14 activity leads to inhibition of  
XX testosterone activity which in turn leads to inhibition of Cyp 4A11  
XX activity. The method of the invention can be used for treating  
XX hypertension in an individual. The molecules of the invention are also  
XX useful for treating disease states associated with lipid metabolism,  
XX pancreatic dysfunction, obesity, type II diabetes, and other  
XX cardiovascular diseases. The present nucleic acid sequence encodes the  
XX human Cyp 4A11 protein of the invention, as described above.  
SQ Sequence 2576 BP; 572 A; 793 C; 583 G; 628 T; 0 other;  
Query Match 100.0%; Score 2576; DB 24; Length 2576;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAATTCGCGAGATCCAGCAGGTGCTGACCATGATGTCTGTGTGAGCCCCAGCA 60  
Db 1 GAATTCGCGAGATCCAGCAGGTGCTGACCATGATGTCTGTGTGAGCCCCAGCA 60  
QY 61 GACTCCTGGGTGATGTCTCTGGAATCTTCAAGCGGCTCCCTGTCTGCTTCTGTC 120  
Db 61 GACTCCTGGGTGATGTCTCTGGAATCTTCAAGCGGCTCCCTGTCTGCTTCTGTC 120  
QY 121 TGCTGATCAAGGAGTTCACTCTACCTGACAGGAGTGTGCTGAAAGCCCTCCAGC 180  
Db 121 TGCTGATCAAGGAGTTCACTCTACCTGACAGGAGTGTGCTGAAAGCCCTCCAGC 180  
QY 181 AGTTCCGCTGCTCTCTCCCACTGCTTTGGGACATCCAGAGCTCCACAAGACC 240  
Db 181 AGTTCCGCTGCTCTCTCCCACTGCTTTGGGACATCCAGAGCTCCACAAGACC 240  
QY 241 AGGAGCTACAACGAGTTCAAAATGGGTGAGACATTCCCAAGTGCCTCTCATTTGGC 300  
Db 241 AGGAGCTACAACGAGTTCAAAATGGGTGAGACATTCCCAAGTGCCTCTCATTTGGC 300  
QY 301 TATGGGAGGCAAGTTCTGTCTCAGCTCTATGACCCTGACTATATGAGGTATTCTGG 360  
Db 301 TATGGGAGGCAAGTTCTGTCTCAGCTCTATGACCCTGACTATATGAGGTATTCTGG 360  
QY 361 GGAGATCAGACCCGAATCCCATGTTCTTCAAGATTCTGCTCAGATGGGTACG 420  
Db 361 GGAGATCAGACCCGAATCCCATGTTCTTCAAGATTCTGCTCAGATGGGTACG 420  
QY 421 GCTTGTCTCTGTTGAATGGGACAGATGTTCCAGCATGACGAGTGTGACCCAGCCT 480



Db 421 GCTGCTCTGTGAAATGGGACAGATGGTTCCAGCATCGACGGATGTCAGCCCGACCT 480  
Qy 481 TCCACTATGACATCTGAAAGCCCTATGTGGGCTCATGGCAGACTCTGTAGAGTATGC 540  
Db 481 TCCACTATGACATCTGAAAGCCCTATGTGGGCTCATGGCAGACTCTGTAGAGTATGC 540  
Qy 541 TGGACAATGGGAAGCTCCTTGGCCAGGATCCCTCTGAGAGTCTTTCAGCAGCTCT 600  
Db 541 TGGACAATGGGAAGCTCCTTGGCCAGGATCCCTCTGAGAGTCTTTCAGCAGCTCT 600  
Qy 601 CCTTGATGACCCCTGACACCATCATGAAGTGTGCTTCAAGCATCAGGGCAGCATCCAGG 660  
Db 601 CCTTGATGACCCCTGACACCATCATGAAGTGTGCTTCAAGCATCAGGGCAGCATCCAGG 660  
Qy 661 TGGACAGGAATTCAGTCTCTACATACAGGCCATTAGTAACTGAAACAACCTGTTTTT 720  
Db 661 TGGACAGGAATTCAGTCTCTACATACAGGCCATTAGTAACTGAAACAACCTGTTTTT 720  
Qy 721 CCCGTGTGAGGAATGCTTTTCAACGAAATGACACCATCTACAGCCTGACCTTGCTGGCC 780  
Db 721 CCCGTGTGAGGAATGCTTTTCAACGAAATGACACCATCTACAGCCTGACCTTGCTGGCC 780  
Qy 781 GCTGGACACACCGCGCTGCGCAGCTGGCCCATCAGACACAGAACCAAGTATCCAACTGA 840  
Db 781 GCTGGACACACCGCGCTGCGCAGCTGGCCCATCAGACACAGAACCAAGTATCCAACTGA 840  
Qy 841 GGAAGGCTCAACTACAGAAAGGAGGAGCTGGAAGAGATCAAGAGGAAGGCAATTTGG 900  
Db 841 GGAAGGCTCAACTACAGAAAGGAGGAGCTGGAAGAGATCAAGAGGAAGGCAATTTGG 900  
Qy 901 ATTTCTGATATCTCTCTTGGCCAAATGAGAAATGGAGCATCTTGTCAACAAGG 960  
Db 901 ATTTCTGATATCTCTCTTGGCCAAATGAGAAATGGAGCATCTTGTCAACAAGG 960  
Qy 961 ACCTCCGTGTGAGGTGACAGTTCATGTTTGAAGGCCACGACACCAAGCAGTGGGA 1020  
Db 961 ACCTCCGTGTGAGGTGACAGTTCATGTTTGAAGGCCACGACACCAAGCAGTGGGA 1020  
Qy 1021 TCTCTGATCTCTATGCTCTGGCCACACACCCCAAGCATCAGAGAGTGCAGGAGG 1080  
Db 1021 TCTCTGATCTCTATGCTCTGGCCACACACCCCAAGCATCAGAGAGTGCAGGAGG 1080  
Qy 1081 AGATCCACAGCCTCTGAGGTGAGAGCTTCCATCAGCTGGAACCACTGGAACAGATGC 1140  
Db 1081 AGATCCACAGCCTCTGAGGTGAGAGCTTCCATCAGCTGGAACCACTGGAACAGATGC 1140  
Qy 1141 CCTACACCAACATGTGATTAAGAGGACATGAGGCTCTAACCCGCTGCCAGGCAATG 1200  
Db 1141 CCTACACCAACATGTGATTAAGAGGACATGAGGCTCTAACCCGCTGCCAGGCAATG 1200  
Qy 1201 GCGAGAGCTCAGACTCCCGTCACTTCCCTGATGGGCGCTCTTGCCCAAGGTATCA 1260  
Db 1201 GCGAGAGCTCAGACTCCCGTCACTTCCCTGATGGGCGCTCTTGCCCAAGGTATCA 1260  
Qy 1261 TGGTCTCTCTCCATTATGCGCTTCAACCAACCCAAAGTGTGSCCAACCAAGAGG 1320  
Db 1261 TGGTCTCTCTCCATTATGCGCTTCAACCAACCCAAAGTGTGSCCAACCAAGAGG 1320  
Qy 1321 TGTTTGACCTTTCCGTTTGGACCGGGTCTGTCTCAACACAGCCAGCTTCTGCCCC 1380  
Db 1321 TGTTTGACCTTTCCGTTTGGACCGGGTCTGTCTCAACACAGCCAGCTTCTGCCCC 1380  
Qy 1381 TCTCAGAGGATCAAGAACTGATTGGGAAAATTTGCAATGAACGAGCTGAAGGTGG 1440  
Db 1381 TCTCAGAGGATCAAGAACTGATTGGGAAAATTTGCAATGAACGAGCTGAAGGTGG 1440  
Qy 1441 CCAGGCGCTGACCTGCTCCGCTTGAAGCTGCTGCTGATCCACAGATCCCATCC 1500  
Db 1441 CCAGGCGCTGACCTGCTCCGCTTGAAGCTGCTGCTGATCCACAGATCCCATCC 1500  
Qy 1501 CCATTGACGACTTGTGTGAATCCAAAATGAATCCACTGCGTCTCAGAGGCTCC 1560  
Db 1501 CCATTGACGACTTGTGTGAATCCAAAATGAATCCACTGCGTCTCAGAGGCTCC 1560

Qy 1561 CTAACCTTGTGAAGCAAGGACAGCTTTGAGGGCTCCACCTGCGCTGCTCTCT 1620  
Db 1561 CTAACCTTGTGAAGCAAGGACAGCTTTGAGGGCTCCACCTGCGCTGCTCTCT 1620  
Qy 1621 GACCCCGCTTCTGTCCCTTCTGCTGTCGCCATATCCTGTTTCTGTCTGCCACCTTC 1680  
Db 1621 GACCCCGCTTCTGTCCCTTCTGCTGTCGCCATATCCTGTTTCTGTCTGCCACCTTC 1680  
Qy 1681 CCTTCTTCCACCTGCTGCTGTCCCGAGTGTGCTGCGCTTCTCTCTCACTTCT 1740  
Db 1681 CCTTCTTCCACCTGCTGCTGTCCCGAGTGTGCTGCGCTTCTCTCTCACTTCT 1740  
Qy 1741 CCAGGCTCCCTACCTGCTGTCTACCTGTCTCTACCCCACTGTATCTTGTGGAGA 1800  
Db 1741 CCAGGCTCCCTACCTGCTGTCTACCTGTCTCTACCCCACTGTATCTTGTGGAGA 1800  
Qy 1801 AAAGCTAGTGTGGAGAAAGCTGAGGCCGAGCTTGCAATGTCTGAATATGTAAAGAG 1860  
Db 1801 AAAGCTAGTGTGGAGAAAGCTGAGGCCGAGCTTGCAATGTCTGAATATGTAAAGAG 1860  
Qy 1861 TCTGAATCATGTCCAGATCCAGGCTTAACCCCTGTGTGCTTGGACCAAGC 1920  
Db 1861 TCTGAATCATGTCCAGATCCAGGCTTAACCCCTGTGTGCTTGGACCAAGC 1920  
Qy 1921 TCTGTGCTGAAGGCTGAAGGCTAACCTGACGCAACATAATCTAAGCCCGGCAATAAA 1980  
Db 1921 TCTGTGCTGAAGGCTGAAGGCTAACCTGACGCAACATAATCTAAGCCCGGCAATAAA 1980  
Qy 1981 CCCCTGTGCTTGGATGAATCCAGGCTGCTGCTTGAATGTGTGACTTGCTG 2040  
Db 1981 CCCCTGTGCTTGGATGAATCCAGGCTGCTGCTTGAATGTGTGACTTGCTG 2040  
Qy 2041 GCTCCTGCTCCTTGTCTTCCAGGATCAATTGTATCTTGAAGTAAAGAACTGCTCTC 2100  
Db 2041 GCTCCTGCTCCTTGTCTTCCAGGATCAATTGTATCTTGAAGTAAAGAACTGCTCTC 2100  
Qy 2101 CATTAATCAAGTAACAGAGCAGATGCTAAACCGTCAAGCTGTAAATGTGTCTTAAT 2160  
Db 2101 CATTAATCAAGTAACAGAGCAGATGCTAAACCGTCAAGCTGTAAATGTGTCTTAAT 2160  
Qy 2161 GCAACATGCCCTTTCGACCCACCCCATCTTCTCAACCACTGTTCTTGTGTGATCACA 2220  
Db 2161 GCAACATGCCCTTTCGACCCACCCCATCTTCTCAACCACTGTTCTTGTGTGATCACA 2220  
Qy 2221 ATAAATPATCTGCACTTCCAGAGCTCGGGGCTTCAAGCCTTCACTAGCTTGGCG 2280  
Db 2221 ATAAATPATCTGCACTTCCAGAGCTCGGGGCTTCAAGCCTTCACTAGCTTGGCG 2280  
Qy 2281 CCCTGACCCCACTTCTCTCAAACTGTCTTCTCACTGCTTGAATCTGCGGACTT 2340  
Db 2281 CCCTGACCCCACTTCTCTCAAACTGTCTTCTCACTGCTTGAATCTGCGGACTT 2340  
Qy 2341 TGTCACCCCCAAGACCTGTGTGGGTCTGAACAACCCCAACATCCCTGAATCTCCACCA 2400  
Db 2341 TGTCACCCCCAAGACCTGTGTGGGTCTGAACAACCCCAACATCCCTGAATCTCCACCA 2400  
Qy 2401 CCTCCAAAACCTGCTGCTGCTTCCAGACTGTCTGCCATACACTGTCTCTTCTCT 2460  
Db 2401 CCTCCAAAACCTGCTGCTGCTTCCAGACTGTCTGCCATACACTGTCTCTTCTCT 2460  
Qy 2461 GCCTGCTGTGTCTGTCTTAATTAATTTCTATTAATCTGCTGTAACTATCACAATC 2520  
Db 2461 GCCTGCTGTGTCTGTCTTAATTAATTTCTATTAATCTGCTGTAACTATCACAATC 2520  
Qy 2521 TCAGTGAATTTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAATTC 2576  
Db 2521 TCAGTGAATTTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGAATTC 2576

XX ABL68591;  
AC  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Kidney cancer related gene sequence SEQ ID NO:6928.  
XX  
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; anti-neoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.  
PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 28-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.  
XX  
PA (AVAL-) AVALON PHARM.  
XX  
PI Young PB, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Sopet DR, Weaver Z;  
XX  
XX WPI; 2002-188264/24.  
PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set -

XX PS Claim 1; SEQ ID 6928; 44p; English.  
XX  
CC The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytosolic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilm's tumour.  
XX  
SQ Sequence 2576 BP; 572 A; 793 C; 583 G; 628 T; 0 other;  
Query Match 100.0%; Score 2576; DB 24; Length 2576;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAATTCGCGAGAGATCCAGCAGGTCCTGACCATGAGTGTCTGTGCTGAGCCCCAGCA 60  
DB 1 GAATTCGCGAGAGATCCAGCAGGTCCTGACCATGAGTGTCTGTGCTGAGCCCCAGCA 60  
QY 61 GACTCCTGGGTGATGTCTCTGGAATCTCCAAGCGGCTCCTGCTCATTTCTGCTTGGC 120  
DB 61 GACTCCTGGGTGATGTCTCTGGAATCTCCAAGCGGCTCCTGCTCATTTCTGCTTGGC 120  
QY 121 TGCTGATCAAGGCGAGTTCAAGTCTCACTGACAGGCGAGTGGCTCTCAAGCCCTCCAGC 180  
DB 121 TGCTGATCAAGGCGAGTTCAAGTCTCACTGACAGGCGAGTGGCTCTCAAGCCCTCCAGC 180  
QY 181 AGTTCGCGTGCCTCCCTCCCACTGGCTCTTGGGCAATCCAGAGCTCCAACAGAGACC 240  
DB 181 AGTTCGCGTGCCTCCCTCCCACTGGCTCTTGGGCAATCCAGAGCTCCAACAGAGACC 240  
QY 241 AGAGCTACAACGGATTCAAGAAATGGGTGAGACATTCGCCAAGTCCCTCATTTGGC 300  
DB 241 AGAGCTACAACGGATTCAAGAAATGGGTGAGACATTCGCCAAGTCCCTCATTTGGC 300  
QY 301 TATGGGGAGGCAAAAGTTCGTGTCCAGCTCTATGACCCCTGACTATATGAAGTGATCTGG 360  
DB 301 TATGGGGAGGCAAAAGTTCGTGTCCAGCTCTATGACCCCTGACTATATGAAGTGATCTGG 360  
QY 361 GGAGATGAGACCCGAATCCCATGTTCTTACAGATTCTGGCTCCATGATTGGGTACG 420  
DB 361 GGAGATGAGACCCGAATCCCATGTTCTTACAGATTCTGGCTCCATGATTGGGTACG 420  
QY 421 GCTTGCTCTGTGATGGGCAACATGTTCCAGCATGACCGATGATGGGACGCT 480  
DB 421 GCTTGCTCTGTGATGGGCAACATGTTCCAGCATGACCGATGATGGGACGCT 480  
QY 481 TCCACTATGACATCTGGAAGCCCTATGTGGGCTCATGGGACAGCTGTGACAGATGATGC 540  
DB 481 TCCACTATGACATCTGGAAGCCCTATGTGGGCTCATGGGACAGCTGTGACAGATGATGC 540  
QY 541 TGGACAAATGGAAGAGCTCTTGCCAGAGATTCCCTCTGAGAGTCTTTACAGCAGCTCT 600  
DB 541 TGGACAAATGGAAGAGCTCTTGCCAGAGATTCCCTCTGAGAGTCTTTACAGCAGCTCT 600  
QY 601 CCTGTGACCTTGACACACCATCAAGAAGTGTGCTTACGCCATCAGGGCAGCATCCAGG 660  
DB 601 CCTGTGACCTTGACACACCATCAAGAAGTGTGCTTACGCCATCAGGGCAGCATCCAGG 660  
QY 661 TGGACGAGAAATTTCACTCTCACTACATACAGGCCATTAGTACTGAACAACCTGTTT 720  
DB 661 TGGACGAGAAATTTCACTCTCACTACATACAGGCCATTAGTACTGAACAACCTGTTT 720



XX 02-OCT-2001, 2001WO-US30589.  
XX  
PR 02-OCT-2000, 2000US-237054P.  
XX  
PA (GENE-) GENE LOGIC INC.  
XX  
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;  
XX  
DR MPI, 2002-426119/45.  
XX  
PT Diagnosing and detecting the progression of liver cancer,  
PT hepatocellular carcinoma or metastatic liver tumor in a patient,  
PT involves detecting the level of expression of two or more genes in a  
PT liver tissue sample -  
XX  
PS Claim 1: SEQ ID NO 1610; 298bp; English.  
XX  
CC The invention relates to a novel method for diagnosing and detecting the  
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver  
CC tumour in a patient, and differentiating metastatic liver cancer from  
CC hepatocellular carcinoma in a patient, involving detecting the level of  
CC expression of two or more genes represented in ABN93503-ABN97455 in a  
CC tissue sample. The method of the invention has hepatotropic, and  
CC cytostatic activity. The method is useful for diagnosing and detecting  
CC the progression of liver cancer, hepatocellular carcinoma and metastatic  
CC liver carcinoma in a patient. The method is useful for identifying  
CC expression profiles which serve as useful diagnostic markers as well as  
CC markers that can be used to monitor disease states, disease progression,  
CC drug toxicity, drug efficacy and drug metabolism.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1763 BP; 381 A; 538 C; 430 G; 414 T; 0 other;  
  
Query Match 63.1%; Score 1625.8; DB 24; Length 1763;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;  
  
QY 10 AGAGATCCAGCAGGTGCTGCACCATGAGTGTCTGTGTGAGCCCGCAGAGACTCTG 69  
DB 1 AGAGATCCAGCAGGTGCTGCACCATGAGTGTCTGTGTGAGCCCGCAGAGACTCTG 60  
  
QY 70 GTGATGTCTTGAATCTCTCAAGCGGCTCCCTGCTCATTTGCTTCTGCTGATCA 129  
DB 61 GTGATGTCTTGAATCTCTCAAGCGGCTCCCTGCTCATTTGCTTCTGCTGATCA 120  
  
QY 130 AGGCAATTCAGCTCTACCTGCACAGGAGTGCTGCTCAAGCCCTCCAGAGTTCCG 189  
DB 121 AGGCAATTCAGCTCTACCTGCACAGGAGTGCTGCTCAAGCCCTCCAGAGTTCCG 180  
  
QY 190 GCCCTCCCTCCCACTGGCTCTTGGGCAATCCAGAGTCTCAACAGACAGAGCTAC 249  
DB 181 GCCCTCCCTCCCACTGGCTCTTGGGCAATCCAGAGTCTCAACAGACAGAGCTAC 240  
  
QY 250 AACGATTCAGAAATGGGTGAGACATCCCAAGTGCTGCTCATTTGGCTATGGGAG 309  
DB 241 AACGATTCAGAAATGGGTGAGACATCCCAAGTGCTGCTCATTTGGCTATGGGAG 300  
  
QY 310 GCAAAATTCGTGCTCAGCTCTATGACCTGACTATATGAAGTGATTTGGGAGATCA 369  
DB 301 GCAAAATTCGTGCTCAGCTCTATGACCTGACTATATGAAGTGATTTGGGAGATCA 360  
  
QY 370 ACCCGAAATCCCATGTTCTCTACAGATTCCTGGCTCCATGATGGGTACGGCTTGC 429  
DB 361 ACCCGAAATCCCATGTTCTCTACAGATTCCTGGCTCCATGATGGGTACGGCTTGC 420  
  
QY 430 TGTGAATGGGAGACATGTTCCAGATTCAGCGATGTCAGCCAGCCTTCCACTATG 489  
DB 421 TGTGAATGGGAGACATGTTCCAGATTCAGCGATGTCAGCCAGCCTTCCACTATG 480  
  
QY 490 ACATCTGAAGCCCTATGTGGGCTCATGGCAGACTCTGTACAGATGATGCTGACAAAT 549

DB 481 ACATCTGAAGCCCTATGTGGGCTCATGGCAGACTCTGTACGAGTGAATGCTGACAAAT 540  
  
QY 550 GGAAGAGCTCTTGGCCAGGATTCCTCTGAGGTCTTTACAGACGTCCTTGATGA 609  
DB 541 GGAAGAGCTCTTGGCCAGGATTCCTCTGAGGTCTTTACAGACGTCCTTGATGA 600  
  
QY 610 CCCTGACACCATCATGAAGTGTGCTTCAAGCCATCAGGGCAGCATCCAGTGACAGGA 669  
DB 601 CCCTGACACCATCATGAAGTGTGCTTCAAGCCATCAGGGCAGCATCCAGTGACAGGA 660  
  
QY 670 ATTCTAGTCTTACATACAGGCCATTAGTGAAGCTGAACCAACCTGTTTCCCGTGA 729  
DB 661 ATTCTAGTCTTACATACAGGCCATTAGTGAAGCTGAACCAACCTGTTTCCCGTGA 720  
  
QY 730 GGAATGCTTTCACCAAGATGACACCATCTACAGCTGACCTGCTGCGCGCTGACAC 789  
DB 721 GGAATGCTTTCACCAAGATGACACCATCTACAGCTGACCTGCTGCGCGCTGACAC 780  
  
QY 790 ACCGCGCTGCGAGCTGCGCCATCAGCAC--AGACCAAGTGAATCCAAGTGAAGAG 846  
DB 781 ACCGCGCTGCGAGCTGCGCCATCAGCACAGAGACCCAAAGTGAATCCAAGTGAAGAG 840  
  
QY 847 CTCACCTACAGAGAGGGGAGCTGAGAGATCAAGAGAGAGGCAATTTGATTTTC 906  
DB 841 CTCACCTACAGAGAGGGGAGCTGAGAGATCAAGAGAGAGGCAATTTGATTTTC 900  
  
QY 907 TGATATCTCTCTCTTGGCCAAATGAGATGGAGCATTTGTACAGCAAGACCTTC 966  
DB 901 TGATATCTCTCTCTTGGCCAAATGAGATGGAGCATTTGTACAGCAAGACCTTC 960  
  
QY 967 GTGCTGAGTGAACAGTTCAATGTTTGAAGGCCACGACACCAAGCAGAGTGGATCTCT 1026  
DB 961 GTGCTGAGTGAACAGTTCAATGTTTGAAGGCCACGACACCAAGCAGAGTGGATCTCT 1020  
  
QY 1027 GGATCTCTATGCTCTGGCCACACACCCAGCATCAGAGAGGTGCGGAGAGATCC 1086  
DB 1021 GGATCTCTATGCTCTGGCCACACACCCAGCATCAGAGAGGTGCGGAGAGATCC 1080  
  
QY 1087 ACAGCCTCTGAGTGAAGACCTCCATCACTGGAACCACTGACCAAGATGCCCTACA 1146  
DB 1081 ACAGCCTCTGAGTGAAGACCTCCATCACTGGAACCACTGACCAAGATGCCCTACA 1140  
  
QY 1147 CCACCATGTGATTAAGAGGCACTGAGGCTTACCACCGGTGCCAGGATTTGGACAG 1206  
DB 1141 CCACCATGTGATTAAGAGGCACTGAGGCTTACCACCGGTGCCAGGATTTGGACAG 1200  
  
QY 1207 AGCTAGCACTCCGCTCACTTCCCTGAATGGGCGCTCCTTGCCCAAGATCATGCTCC 1266  
DB 1201 AGCTAGCACTCCGCTCACTTCCCTGAATGGGCGCTCCTTGCCCAAGATCATGCTCC 1260  
  
QY 1267 TCCTCTCATTTATGGCCTTCAACCAACCAAAAGTGTGGCCCAAGAGGTGTTG 1326  
DB 1261 TCCTCTCATTTATGGCCTTCAACCAACCAAAAGTGTGGCCCAAGAGGTGTTG 1320  
  
QY 1327 ACCCTTCCGTTTTCACCGGGTTCGCTCAACACAGCAGCTTCTCTCTCTCTCTCTCT 1386  
DB 1321 ACCCTTCCGTTTTCACCGGGTTCGCTCAACACAGCAGCTTCTCTCTCTCTCTCTCTCT 1380  
  
QY 1387 GAGGATCAAGAACTGCATTTGGGAAACAATTTGCCATGAACGAGCTGAAGGTGGCC--A 1443  
DB 1381 GAGGATCAAGAACTGCATTTGGGAAACAATTTGCCATGAACGAGCTGAAGGTGGCCAGC 1440  
  
QY 1444 CGGCGCTGACCTGCTCCGCTTGAAGTGTGCTGCTGATCCCAACAGATCCCATCCCA 1503  
DB 1441 AGGCGCTGACCTGCTCCGCTTGAAGTGTGCTGCTGATCCCAACAGATCCCATCCCA 1500  
  
QY 1504 TTGACAGACTTGTGAATCCAAAATGGAATCCACTGCGTCTCAGAGGCTCCCTA 1563  
DB 1501 TTGACAGACTTGTGAATCCAAAATGGAATCCACTGCGTCTCAGAGGCTCCCTA 1560  
  
QY 1564 ACCCTTGTGAAGACAAAGACCAAGCTTTGAAGGCTCCACTGCGCTCTCTCTCTGAC 1623





QY	490	ACATCCTGAAGCCCTTATGTGGGGCTCATGGCAGACTCTGTACGAGTGATGCTGGACAAT	549
Db	481	ACATCCTGAAGCCCTATGTGGGGCTCATGGCAGACTCTGTACGAGTGATGCTGGACAAT	540
QY	550	GGGAAGAGCTCCTTGGCCAGGATCCCTCTGGAGGCTTTAGCAGCTCTCTGTATGA	609
Db	541	GGGAAGAGCTCCTTGGCCAGGATCCCTCTGGAGGCTTTAGCAGCTCTCTGTATGA	600
QY	610	CCCTGGACACCATCATGAAGTGTGCTTCAAGCATCAGGGCAGCATCCAGGTGGACAGA	669
Db	601	CCCTGGACACCATCATGAAGTGTGCTTCAAGCATCAGGGCAGCATCCAGGTGGACAGA	660
QY	670	ATTCTAGTCTTACATACAGGCCATTAGTGAACCTGAACAACCTGTTTTTCCCGTGA	729
Db	661	ATTCTAGTCTTACATACAGGCCATTAGTGAACCTGAACAACCTGTTTTTCCCGTGA	720
QY	730	GGAATGCCCTTTCACCAAGATGACCATCTACAGCCTGACCTCTGCTGGCCGCTGGAC	789
Db	721	GGAATGCCCTTTCACCAAGATGACCATCTACAGCCTGACCTCTGCTGGCCGCTGGAC	780
QY	790	ACCGGCCCTGCCAGCTGGCCCATCAGCACAC--AGACCAAGTATCCAACTGAGGAAG	846
Db	781	ACCGGCCCTGCCAGCTGTCTCCATCAGCACACGAGACCCAAAGTATCCAACTGAGGAAG	840
QY	847	CTCAACTACAGAGGAGGGGAGCTGAGAGATCAAGAGAGAGGCAATTGGATTTTC	906
Db	841	CTCAACTACAGAGGAGGGGAGCTGAGAGATCAAGAGAGAGGCAATTGGATTTTC	900
QY	907	TGATATCCTCCTCTTGGCCAAATGAGATGGAGCACTTGTCAAGACAAGACCTCC	966
Db	901	TGATATCCTCCTCTTGGCCAAATGAGATGGAGCACTTGTCAAGACAAGACCTCC	960
QY	967	GTGCTGAGGTGACACGTTCACTTTGAGGGCCACGACACACAGCCAGTGGATCTCT	1026
Db	961	GTGCTGAGGTGACACGTTCACTTTGAGGGCCACGACACACAGCCAGTGGATCTCT	1020
QY	1027	GGATCCTCTATGCTCTGGCCACACACCCCAAGCATCAGAGAGGTGCCGGAGAGATCC	1086
Db	1021	GGATCCTCTATGCTCTGGCCACACACCCCAAGCATCAGAGAGGTGCCGGAGAGATCC	1080
QY	1087	ACAGCCTCCTGGGTGATGAGCCTTCATCACCTGGAACCACTGGAACAGATGCCCTACA	1146
Db	1081	ACAGCCTCCTGGGTGATGAGCCTTCATCACCTGGAACCACTGGAACAGATGCCCTACA	1140
QY	1147	CCACCATGTGCATTAAAGAGGCACTGAGGCTCTACCCACCGGTGCCAGGCAATTGGCAG	1206
Db	1141	CCACCATGTGCATTAAAGAGGCACTGAGGCTCTACCCACCGGTGCCAGGCAATTGGCAG	1200
QY	1207	AGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCTTGCCCAAGGTATCATGTGCC	1266
Db	1201	AGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCTTGCCCAAGGTATCATGTGCC	1260
QY	1267	TCCTCTCATTATGAGCCTTCAACAACCAAAAGTGGCCCAACCCAGAGGTGTTG	1326
Db	1261	TCCTCTCATTATGAGCCTTCAACAACCAAAAGTGGCCCAACCCAGAGGTGTTG	1320
QY	1327	ACCCTTTCGTTTGCACCGGGTTCTGCTCAACAAGCCCTTCTGCTGCCCTTCTAG	1386
Db	1321	ACCCTTTCGTTTGCACCGGGTTCTGCTCAACAAGCCCTTCTGCTGCCCTTCTAG	1380
QY	1387	GAGGATCAAGGAACCTGCATTGGGAAACAATTGGCCATGAAGAGCTGAAGGTGGCC--A	1443
Db	1381	GAGGATCAAGGAACCTGCATTGGGAAACAATTGGCCATGAAGAGCTGAAGGTGGCCAGC	1440
QY	1444	CGGCTCTGACCTGCTCCGCTTGAAGTGTGCTGCTGATCCACCAAGATCCCATCCCA	1503
Db	1441	AGGCTCTGACCTGCTCGCTTGAAGTGTGCTGCTGATCCACCAAGATCCCATCCCA	1500
QY	1504	TTGCAAGCACTTGTGTTGAATCCAAATAATGGAATCCACTGCGCTCAGAGAGGCTCCCTA	1563
Db	1501	TTGCAAGCACTTGTGTTGAATCCAAATAATGGAATCCACTGCGCTCAGAGAGGCTCCCTA	1560
QY	1564	ACCCTTGTGAAGCAAGAGCAAGCTTTGAGGGCTCCACTGCGCTCTGTCTTCTGTAC	1623

Db	Accession	Sequence	Length
Db	1561	ACCGTTGTGAAGACAAGACCAAGCCTTGTAGGGGCTCCACCTGCGGCTGTCTTCTGAC	1620
OY	1624	CCCCGCTTGTGTCCTTCTGTCGCCATATCTGTTTCTGTGTGTCGCCACCTTCCCT	1683
Db	1621	CCCCGCTTGTGTCCTTCTGTCGCCATATCTGTTTCTGTGTGTCGCCATACACCTG	1680
OY	1684	TCTTC	1688
Db	1681	TCTCC	1685
RESULT 6			
ID	ABL68599	standard; DNA; 1763 BP.	
AC	ABL68599;		
DT	15-MAY-2002	(first entry)	
DE	Kidney cancer related gene sequence SEQ ID NO:6936.		
KW	Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;		
KW	stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;		
KW	cytostatic; gene therapy; antineoplastic; wilm's tumour; adenocarcinoma;		
KW	gene; ds.		
OS	Homo sapiens.		
PN	WO200194629-A2.		
PD	13-DEC-2001.		
PF	30-MAY-2001; 2001WO-US10838.		
PR	05-JUN-2000; 2000US-209473P.		
PR	05-JUN-2000; 2000US-209531P.		
PR	18-SEP-2000; 2000US-233133P.		
PR	18-SEP-2000; 2000US-233617P.		
PR	20-SEP-2000; 2000US-234009P.		
PR	20-SEP-2000; 2000US-234034P.		
PR	20-SEP-2000; 2000US-234052P.		
PR	22-SEP-2000; 2000US-234509P.		
PR	22-SEP-2000; 2000US-234567P.		
PR	25-SEP-2000; 2000US-234923P.		
PR	25-SEP-2000; 2000US-234924P.		
PR	25-SEP-2000; 2000US-235077P.		
PR	25-SEP-2000; 2000US-235082P.		
PR	25-SEP-2000; 2000US-235134P.		
PR	25-SEP-2000; 2000US-235280P.		
PR	26-SEP-2000; 2000US-235637P.		
PR	26-SEP-2000; 2000US-235638P.		
PR	27-SEP-2000; 2000US-235711P.		
PR	27-SEP-2000; 2000US-235720P.		
PR	27-SEP-2000; 2000US-235840P.		
PR	27-SEP-2000; 2000US-235863P.		
PR	28-SEP-2000; 2000US-236028P.		
PR	28-SEP-2000; 2000US-236032P.		
PR	28-SEP-2000; 2000US-236033P.		
PR	28-SEP-2000; 2000US-236034P.		
PR	28-SEP-2000; 2000US-236109P.		
PR	28-SEP-2000; 2000US-236111P.		
PR	29-SEP-2000; 2000US-236842P.		
PR	29-SEP-2000; 2000US-236891P.		
PR	02-OCT-2000; 2000US-237172P.		
PR	02-OCT-2000; 2000US-237173P.		
PR	02-OCT-2000; 2000US-237278P.		
PR	02-OCT-2000; 2000US-237294P.		
PR	02-OCT-2000; 2000US-237316P.		
PR	03-OCT-2000; 2000US-237425P.		
PR	03-OCT-2000; 2000US-237598P.		
PR	03-OCT-2000; 2000US-237604P.		

PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.

XX (AVAL-) AVALON PHARM.

XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppel DR, Weaver Z;

XX WPI; 2002-188264/24.

PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set -

XX Claim 1; SEQ ID 6936; 44pp; English.

XX The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilms' tumour.

XX Sequence 1763 BP; 381 A; 538 C; 430 G; 414 T; 0 other;

Query Match 63.1%; Score 1625.8; DB 24; Length 1763;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;

QY 10 AGAGATCCAGCAGGTGCTGCACCATGAGTGTCTGTGTGAGCCCAAGCACTCTGG 69  
DB 1 AGAGATCCAGCAGGTGCTGCACCATGAGTGTCTGTGTGAGCCCAAGCACTCTGG 60  
QY 70 GTGATGTCTTGAATCTCTCAAGCGGCTCCCTGCTCATTTCTGCTTCTGCTGATCA 129  
DB 61 GTGATGTCTTGAATCTCTCAAGCGGCTCCCTGCTCATTTCTGCTTCTGCTGATCA 120  
QY 130 AGGCAATTGAGTCTTACTGTCAGAGGAGTGGTCTCAAGCCCTCCAGAGTTCCCGT 189  
DB 121 AGGCAATTGAGTCTTACTGTCAGAGGAGTGGTCTCAAGCCCTCCAGAGTTCCCGT 180  
QY 190 GCCCTCCCTCCCACTGGCTCTTGGGACATCCAGAGCTCCAAACAGACAGAGCTAC 249  
DB 181 GCCCTCCCTCCCACTGGCTCTTGGGACATCCAGAGCTCCAAACAGACAGAGCTAC 240  
QY 250 AACGATTGAGAAATGGGTGAGACATTCACCAAGTCTGCTCATTTGGCTATGGGAG 309  
DB 241 AACGATTGAGAAATGGGTGAGACATTCACCAAGTCTGCTCATTTGGCTATGGGAG 300  
QY 310 GCAAGTTGCTGTCCAGCTCTATGACCTGACTATATGAAGGTGATTCTGGGAGATCAG 369  
DB 301 GCAAGTTGCTGTCCAGCTCTATGACCTGACTATATGAAGGTGATTCTGGGAGATCAG 360  
QY 370 ACCCGAATCCCATGTTCTTACAGATTCTGCTGCTCATGATTGGGTACGGCTTCTCC 429  
DB 361 ACCCGAATCCCATGTTCTTACAGATTCTGCTGCTCATGATTGGGTACGGCTTCTCC 420  
QY 430 TGTGAATGGGACAGATGTTCTCAGCATGACGAGTGTGACCCAGCCTTCCACTATG 489  
DB 421 TGTGAATGGGACAGATGTTCTCAGCATGACGAGTGTGACCCAGCCTTCCACTATG 480

QY 490 ACATCTGAAGCCCTATGTGGGCTCATGCGACACTCTGTACGAGTGTGACCAAT 549  
DB 481 ACATCTGAAGCCCTATGTGGGCTCATGCGACACTCTGTACGAGTGTGACCAAT 540  
QY 550 GGAAGAGCTCTTGGCAGAGTATCCCTCTGAGGTCTTTCAGACAGTCTCTGATGA 609  
DB 541 GGAAGAGCTCTTGGCAGAGTATCCCTCTGAGGTCTTTCAGACAGTCTCTGATGA 600  
QY 610 CCTGACACCATCATGAAAGTGTGCTTCAAGCATCAGGCGAGCATCCAGGTGACAGA 669  
DB 601 CCTGACACCATCATGAAAGTGTGCTTCAAGCATCAGGCGAGCATCCAGGTGACAGA 660  
QY 670 ATTCTAGTCTCATACAGGCCATTAGTACCTGAAACAACCTGTTTTCCTGCTGA 729  
DB 661 ATTCTAGTCTCATACAGGCCATTAGTACCTGAAACAACCTGTTTTCCTGCTGA 720  
QY 730 GGAATGCTTTACACAGATGACACCATCTACAGCTGACCTCTGCTGGCGCTGACAC 789  
DB 721 GGAATGCTTTACACAGATGACACCATCTACAGCTGACCTCTGCTGGCGCTGACAC 780  
QY 790 ACCGCGCTGACAGTGGCCCATCAGACACACACACACACACACACACACACACAC 846  
DB 781 ACCGCGCTGACAGTGGCCCATCAGACACACACACACACACACACACACACACAC 840  
QY 847 CTCACTACAGAAAGAGGAGAGTGAAGATCAAGAGAGAGAGAGAGAGAGAGAGAG 906  
DB 841 CTCACTACAGAAAGAGGAGAGTGAAGATCAAGAGAGAGAGAGAGAGAGAGAGAG 900  
QY 907 TGGATATCTCTCTTGGCCAAATGAGAGATGGAGCATTTGTACAGACAGACCTCC 966  
DB 901 TGGATATCTCTCTTGGCCAAATGAGAGATGGAGCATTTGTACAGACAGACCTCC 960  
QY 967 GTGCTGAGTGAACAGCTTCATGTTGAGGGCCACAGACACAGCCAGTGGGATCTCT 1026  
DB 961 GTGCTGAGTGAACAGCTTCATGTTGAGGGCCACAGACACAGCCAGTGGGATCTCT 1020  
QY 1027 GGAATCTCTATGCTTGGCCACACACCCCAAGCATCAGAGAGTGGCGGAGAGATCC 1086  
DB 1021 GGAATCTCTATGCTTGGCCACACACCCCAAGCATCAGAGAGTGGCGGAGAGATCC 1080  
QY 1087 ACAGCCTCTGCTGATGAGAGCTCTCATCAGTGAACACCTGAGACAGATGCTTACA 1146  
DB 1081 ACAGCCTCTGCTGATGAGAGCTCTCATCAGTGAACACCTGAGACAGATGCTTACA 1140  
QY 1147 CCACTATGTGATTAAGAGGCACTGAGGCTTACCCACCGGTGCAAGCATTGGCAGAG 1206  
DB 1141 CCACTATGTGATTAAGAGGCACTGAGGCTTACCCACCGGTGCAAGCATTGGCAGAG 1200  
QY 1207 AGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCTTCCCAAAAGTATCAATGCTCC 1266  
DB 1201 AGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCTTCCCAAAAGTATCAATGCTCC 1260  
QY 1267 TCCCTCTCATTTATGAGCTTCAACCAACCAAAAGTGTGAGCCCAACCAAGGTGTTG 1326  
DB 1261 TCCCTCTCATTTATGAGCTTCAACCAACCAAAAGTGTGAGCCCAACCAAGGTGTTG 1320  
QY 1327 ACCCTTCCGTTTGAACCGGGTCTGCTCAACACAGCCAGCTTCTCTGCTCTCAG 1386  
DB 1321 ACCCTTCCGTTTGAACCGGGTCTGCTCAACACAGCCAGCTTCTCTGCTCTCAG 1380  
QY 1387 GAGATCAAGGAAGTGTGAGGAAACAATTTGCCATGAAAGAGTGAAGGTGGCC--A 1443  
DB 1381 GAGATCAAGGAAGTGTGAGGAAACAATTTGCCATGAAAGAGTGAAGGTGGCCAGC 1440  
QY 1444 CGGCGCTGACCTGCTCGCTTGAAGCTGCTGATCCCAACAGATCCCAATCCCA 1503  
DB 1441 AGGCGCTGACCTGCTCGCTTGAAGCTGCTGATCCCAACAGATCCCAATCCCA 1500  
QY 1504 TTGCAAGACTTGTGTAATCCAAAATGAAATCCACTGCGTCTCAGAGGCTCCCTA 1563  
DB 1501 TTGCAAGACTTGTGTAATCCAAAATGAAATCCACTGCGTCTCAGAGGCTCCCTA 1560

OY 1564 ACCCTGTGAGACAAGACAGCCTTTGAGGGCCTCCACCTGCCGTCTGTCTTCTTGAC 1623  
DB 1561 ACCCTGTGAGACAAGACAGCCTTTGAGGGCCTCCACCTGCCGTCTGTCTTCTTGAC 1620  
OY 1624 CCCCCGCTTGTGTCCCTTCTGTCTGCCCCATATCTCTGTTTCTGTCTGCCACCTTCCCT 1683  
DB 1621 CCCCCGCTTGTGTCCCTTCTGTCTGCCCCATATCTCTGTTTCTGTCTGCCCATACACCTTG 1680  
OY 1684 TCTTC 1688  
DB 1681 TCTCC 1685

RESULT 7  
ABL68881  
ID ABL68881 standard; DNA; 1763 BP.  
XX  
AC ABL68881;  
XX  
DT 15-MAY-2002 (first entry)  
XX  
DE Kidney cancer related gene sequence SEQ ID NO:7218.

XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KM stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KM cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX

PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.  
PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 28-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.

PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.  
XX  
PA (AVAL-) AVALON PHARM.  
XX  
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX  
DR WPI, 2002-188264/24.  
XX  
PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set  
XX  
PS Claim 1; SEQ ID 7218; 44pp; English.  
XX  
CC The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilm's tumour.  
XX  
SQ Sequence 1763 BP; 381 A; 538 C; 430 G; 414 T; 0 other;

Query Match 63.1%; Score 1625.8; DB 24; Length 1763;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 1662; Conservative 0; Mismatches 17; Indels 6; Gaps 2;

OY 10 AGAGATCAGCAGGTGCTGCACCATGAGTGTCTGTGTGTCGAGCCCAAGACTCCTGG 69  
DB 1 AGAGATCAGCAGGTGCTGCACCATGAGTGTCTGTGTGTCGAGCCCAAGACTCCTGG 60  
OY 70 GTGATGCTCTGGAATCCTCCAGGGCCTCCCTGCTCATTCGTTGCTGTGATCA 129  
DB 61 GTGATGCTCTGGAATCCTCCAGGGCCTCCCTGCTCATTCGTTGCTGTGATCA 120  
OY 130 AGCAGTTCACTCTACCTGCACAGCAGTGGCTCTCAAGCCCTCCAGAGTTCCCGT 189  
DB 121 AGCAGTTCACTCTACCTGCACAGCAGTGGCTCTCAAGCCCTCCAGAGTTCCCGT 180  
OY 190 GCCCTCCTCCCACTGGCTCTTCGGGCAATCCAGAGGCTCCACAGACAGAGCTTAC 249  
DB 181 GCCCTCCTCCCACTGGCTCTTCGGGCAATCCAGAGGCTCCACAGACAGAGCTTAC 240  
OY 250 AACGATTCAGAAATGGGTGAGACATTCACAGTGCCTGCTCATTTGGCTATGGGAG 309  
DB 241 AACGATTCAGAAATGGGTGAGACATTCACAGTGCCTGCTCATTTGGCTATGGGAG 300  
OY 310 GCAAAGTCGTGTCCAGCTCTATGACCTGACTATATGAAGGTGATTCGGAGATCAG 369  
DB 301 GCAAAGTCGTGTCCAGCTCTATGACCTGACTATATGAAGGTGATTCGGAGATCAG 360  
OY 370 ACCCGAATCCCATGTGTTCTTACAGATTCTGGCTCCATGATGGGTACGGCTTCTCC 429  
DB 361 ACCCGAATCCCATGTGTTCTTACAGATTCTGGCTCCATGATGGGTACGGCTTCTCC 420  
OY 430 TGTGAAATGGGCAAGACATGTGTTCCAGCATGACGAGATGCTGACCCCAAGCTTCCACTATG 489



```

Db      421 TGTGAATGGGCAAGACATGGTTCAGCATCGACGGATGCTGACCCCAAGCTTCCACTATG 480
QY      490 ACATCCTGAAGCCCTATATGTGGGCTCATGGCAGACTCTGTACGAGTGTCTGGACAAT 549
Db      481 ACATCCTGAAGCCCTATATGTGGGCTCATGGCAGACTCTGTACGAGTGTCTGGACAAT 540
QY      550 GGGAGAAGCTCCTTGGCCAGGATTCCTCTGAGAGTCTTTCAGCAGCTCTCTTGATGA 609
Db      541 GGGAGAAGCTCCTTGGCCAGGATTCCTCTGAGAGTCTTTCAGCAGCTCTCTTGATGA 600
QY      610 CCCTGACACCATCATGAAGTGTGCTTCAGCCATCAGGCGAGCATCCAGGTGACAGA 669
Db      601 CCCTGACACCATCATGAAGTGTGCTTCAGCCATCAGGCGAGCATCCAGGTGACAGA 660
QY      670 ATTCTCAGTCTCATATACAGGCCATTAGTGAACCTGTTTTCCTGTTGA 729
Db      661 ATTCTCAGTCTCATATACAGGCCATTAGTGAACCTGTTTTCCTGTTGA 720
QY      730 GGAATGCTTTCACCAAGATGACACCATCTACAGCTGACCTGCTGCGCTGACAC 789
Db      721 GGAATGCTTTCACCAAGATGACACCATCTACAGCTGACCTGCTGCGCTGACAC 780
QY      790 ACCGCGCTGCGCAGCTGGCCCATCAGCACAC---AGACCAAGTATCCACTGAGGAAG 846
Db      781 ACCGCGCTGCGCAGCTGGCCCATCAGCACACCAAGTATCCACTGAGGAAG 840
QY      847 CTCAACTACAGAAGAGGGGAGCTGAGAGAATCAAGAGAAGAGGATTGATTTC 906
Db      841 CTCAACTACAGAAGAGGGGAGCTGAGAGAATCAAGAGAAGAGGATTGATTTC 900
QY      907 TGGATATCTCTCTTGGCCAAATGAGAAATGGAGCATCTTGTCAACAAGACCTCC 966
Db      901 TGGATATCTCTCTTGGCCAAATGAGAAATGGAGCATCTTGTCAACAAGACCTCC 960
QY      967 GTGCTGAGTGACACGTTTCATGTTGAGGGCCACGACACCAAGCCAGTGGGATCTCT 1026
Db      961 GTGCTGAGTGACACGTTTCATGTTGAGGGCCACGACACCAAGCCAGTGGGATCTCT 1020
QY      1027 GGATCCTCTATGCTGTGGCCACACCCCAAGCATCAGAGAGTGGCGGAGAGATCC 1086
Db      1021 GGATCCTCTATGCTGTGGCCACACCCCAAGCATCAGAGAGTGGCGGAGAGATCC 1080
QY      1087 ACAGCCTCTGGTGAATGAGCCTTCATCACCTGGAACCACTGACCAATGCCCTACA 1146
Db      1081 ACAGCCTCTGGTGAATGAGCCTTCATCACCTGGAACCACTGACCAATGCCCTACA 1140
QY      1147 CCAACCATGTGATTAAAGAGGCACTGAGGCTTACCCACCGGTGCCAGGATGGCAGAG 1206
Db      1141 CCAACCATGTGATTAAAGAGGCACTGAGGCTTACCCACCGGTGCCAGGATGGCAGAG 1200
QY      1207 AGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCTTCCCAAGGTAATCATGTGCC 1266
Db      1201 AGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCTTCCCAAGGTAATCATGTGCC 1260
QY      1267 TCCTCTCATTTATAGGCTTCAACAACCCAAAGTGTGGCCCAACCCAGAGGTGTTG 1326
Db      1261 TCCTCTCATTTATAGGCTTCAACAACCCAAAGTGTGGCCCAACCCAGAGGTGTTG 1320
QY      1327 ACCCTTTCGTTTTCACACCGGGTCTGCTCAACACAGCCAGCTTTCCTGCTTCTCAG 1386
Db      1321 ACCCTTTCGTTTTCACACCGGGTCTGCTCAACACAGCCAGCTTTCCTGCTTCTCAG 1380
QY      1387 GAGGATCAAGGAAGTGCATTGGGAAAACAATTGGCCATGAACGAGCTGAAGGTGGCC--A 1443
Db      1381 GAGGATCAAGGAAGTGCATTGGGAAAACAATTGGCCATGAACGAGCTGAAGGTGGCCAGC 1440
QY      1444 CGGCTCTGACCTGCTCCGCTTGGAGCTGCTGCTGATCCCAACGAGATCCCATCCCA 1503
Db      1441 AGGCCCTGACCTCGTCCGCTTGGAGCTGCTGCTGATCCCAACGAGATCCCATCCCA 1500
QY      1504 TTGACGAGCTTGTGTTAAATCAAAATGGAATCCAGTGCCTCAGAGAGCTCCCTA 1563
Db      1501 TTGACGAGCTTGTGTTAAATCAAAATGGAATCCAGTGCCTCAGAGAGCTCCCTA 1560

```

```

QY      1564 ACCCTTGAAGACAAGACAGCAGCTTGAAGGCTCCACCTGCGCTGCTCTGAC 1623
Db      1561 ACCCTTGAAGACAAGACAGCAGCTTGAAGGCTCCACCTGCGCTGCTCTGAC 1620
QY      1624 CCGCGCTTGTCCCTCTCTGCTGACCAATATCTGTTTCTGTGCCACCTTCCCT 1683
Db      1621 CCGCGCTTGTCCCTCTCTGCTGACCAATATCTGTTTCTGTGCCATACACCTG 1680
QY      1684 TCTTC 1688
Db      1681 TCTCC 1685

```

```

RESULT 8
ABK50070
ID      ABK50070 standard; DNA; 1872 BP.
XX
AC      ABK50070;
XX
DT      15-JUL-2002 (first entry)
XX
DE      DNA encoding human Cyp 4A22 protein.
XX
KW      Human; Cyp 4A14; hypertension; Cyp 4A11; testosterone inhibitor; obesity;
KW      lipid metabolism disease; pancreatic dysfunction; type II diabetes;
KW      cardiovascular disease; Cyp 4A22; gene; ds.
XX
OS      Chimeric - Homo sapiens.
OS      Chimeric - Synthetic.
XX
FH      Key
FT      CDS
FT      Location/Qualifiers
FT      /tag= a
FT      /product= "Human Cyp 4A22 protein"
XX
PN      W0200217856-A2.
XX
PD      07-MAR-2002.
XX
PF      29-AUG-2001; 2001WO-US26914.
XX
PR      29-AUG-2000; 2000US-228947P.
XX
PA      (UYVA-) UNIV VANDERBILT.
XX
PI      Capdevila J, Waterman M, Holla V;
XX
XX      WPI; 2002-382929/41.
DR      P-PSDB; AAU79995.
XX
PT      Treating hypertension in an individual by inhibiting testosterone
PT      activity, enhancing Cyp 4A14 activity, or by inhibiting Cyp 4A11
PT      activity in the individual -
XX
PS      Examples; Page 73-74; 80pp; English.
XX
CC      The present invention relates to a new method of treating hypertension in
CC      an individual. The method comprises inhibiting testosterone activity,
CC      enhancing Cyp 4A14 activity, or inhibiting Cyp 4A11 activity in the
CC      individual. Enhancing Cyp 4A14 activity leads to inhibition of
CC      testosterone activity which in turn leads to inhibition of Cyp 4A11
CC      activity. The method of the invention can be used for treating
CC      hypertension in an individual. The molecules of the invention are also
CC      useful for treating disease states associated with lipid metabolism,
CC      pancreatic dysfunction, obesity, type II diabetes, and other
CC      cardiovascular diseases. The present nucleic acid sequence encodes the
CC      human Cyp 4A22 protein of the invention, as described above.
XX
SQ      Sequence 1872 BP; 427 A; 554 C; 481 G; 410 T; 0 other;

```

Query Match 59.2%; Score 1525.2; DB 24; Length 1872;  
 Best Local Similarity 97.6%; Pred. No. 0;

Matches 1548; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

```
OY 7 CGCAGAGATCCAGCAGGTGCTGCACCATGATGTCTCTGTGCTGAGCCCCAGACACTCC 66
DB 287 CTCAGAGATCCAGCAGGTGCTGCACCATGATGTCTCTGTGCTGAGCCCCAGACAGCGCC 346
OY 67 TGGGTGATGTCTCTGGAATCTCCAAAGCGGCTCCCTGCTCATTTCTGCTTCTGTGCTGA 126
DB 347 TGGGTGATGTCTCTGGAATCTCCAAAGTGAATCTCCCTGCTCATTTCTGCTTCTGTGCTGA 406
OY 127 TCAGGCGAGTTCAAGCTCTACCTGCACAGGAGTGCTGCTCAAAAGCCCTCCAGCAGTTCC 186
DB 407 TCAGGCGAGTTCAAGCTCTACCTGCATAGGAGTGCTGCTCAAAAGCCCTCCAGCAGTTCC 466
OY 187 CGTGCCTCCCTCCCACTGGCTCTTGGGCAATCCAGAGCTCCAACAGACAGAGAGC 246
DB 467 CGTGCCTCCCTCCCACTGGCTCTTGGGCAATCCAGAGTTCCAACAGACAGAGAGC 526
OY 247 TACAACGGAATCAGAAATGGGTGAGACATTTCCCAAGTGCCTGTCTCATTTGGCTATGGG 306
DB 527 TACAACGGAATCAGAAACGGGTGAAGACATTTCCCAAGTGCCTGTCTCATTTGAATATGGG 586
OY 307 GAGGCAAGTTGCTGTCCAGCTCTATGACCTGACTATATGAAGGTGATTTCTGGGAGAT 366
DB 587 GAGGCAAGTTGCTGTCCAGCTCTATGACCTGACTATATGAAGGTGATTTCTGGGAGAT 646
OY 367 CAGACCCGAAATCCCATGTTCTCTACAGATTCCTGGCTCCATGAGTTGGGTACGCGCTTG 426
DB 647 CAGACCCGAAATCCCATGTTCTCTACAGATTCCTGGCTCCATGAGTTGGGTACGCGCTTG 706
OY 427 TCCTGTTGAATGGGCAAGCATGTTCCAGCATCGACGGATGCTGACCCCAAGCTTCCACT 486
DB 707 TCCTGTTGAATGGGCAAGCATGTTCCAGCATCGACGGATGCTGACCCCAAGCTTCCACT 766
OY 487 ATGACATCTCTGAAGCCCTATGTGGGCTCATGGCAGACTCTGTAGAGATGATGCTGACA 546
DB 767 ATGACATCTCTGAAGCCATATGTGGGCTCATGGCAGACTCTGTAGAGATGATGCTGACA 826
OY 547 AATGGGAAGAGCTCTTGCCAGAGATTCCTCTGAGGTCTTTCAAGCAGTCTCTTGA 606
DB 827 AATGGGAAGAGCTCTTGCCAGAGATTCCTCTGAGGTCTTTCAAGCAGTCTCTTGA 886
OY 607 TGACCTTGACACCATCATGAAGTGTGCTTCAAGCATCAGGCGACATCCAGGTGACA 666
DB 887 TGACCTTGACACCATCATGAAGTGTGCTTCAAGCATCAGGCGACATCCAGGTGACA 946
OY 667 GGAATTCTCAGTCTCTACATACAGGCGCATTAAGCTGAACAACCTGTTTTTCCCGTG 726
DB 947 GGAATTCTCAGTCTCTACATACAGGCGCATTAAGCTGAACAACCTGTTTTTCTGTGA 1006
OY 727 TGAGGAATGCTTTTCAACGAATGACACCATCTACAGCTGACCTTGTCTGCGCGTGA 786
DB 1007 TGAGGAATGCTTTTCAAGGAATGACACCATCTACAGCTGACCTTGTCTGCGCGTGA 1066
OY 787 CACACCGCGCTGCGAGCTGCGCCATCAGACACAGACCAAGTGAATCCAATGAGGAAG 846
DB 1067 CACACCGCGCTGCGAGCTGCGCCATCAGACACAGACCAAGTGAATCCAATGAGGAAG 1126
OY 847 CTCAACTACAGAGAGGGGAGCTGAGGAAGATCAAGAGGAAGGCAATTTGATTTTC 906
DB 1127 CTCAACTACAGAGAGGGGAGCTGAGGAAGATCAAGAGGAAGGCAATTTGATTTTC 1186
OY 907 TGGATATCTCTCTTGGCCAAATGAGAAATGGAGCATTTGTCAACAAGACCTTC 966
DB 1187 TGGACATCTCTCTTGGCCAAATGAGAAATGGAGCATTTGTCAACAAGACCTTC 1246
OY 967 GTGCTGAGGTGACAGCTTCAATGTTTGAAGGCGCACGACACCAAGCAGTGGATCTCT 1026
DB 1247 GTGCTGAGGTGACAGCTTCAATGTTTGAAGGCGCACGACACCAAGCAGTGGATCTCT 1306
OY 1027 GGATCTCTATGCTCTGGCCACACACCCCAAGCATCAGAGAGGTGCGGAGAGAGATCC 1086
DB 1307 GGATCTCTATGCTCTGGCCACACACCCCAAGCATCAGAGAGGTGCGGAGAGAGATCC 1366
```

```
OY 1087 ACAGCCTCTGGGTGATGAGACCTTCATCACTGGAACCACTGAGACCAATGCCCTACA 1146
DB 1367 ATGGCCTCTGGGTGATGAGACCTTCATCACTGGAACCACTGAGACCAATGCCCTACA 1426
OY 1147 CCACCATGTGATTAAGAGGCACTGAGGCTTAACCAACCGGTGCCAGCATGGCAGAG 1206
DB 1427 CCACCATGTGATTAAGAGGCACTGAGGCTTAACCAACCGGTGCCAGCATGGGAAG 1486
OY 1207 AGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCTTGCCCAAGATCATGTCTC 1266
DB 1487 AGCTCAGCACTCCCGTCACTTCCCTGATGGGCGCTCTTGCCCAAGATCATGTCTC 1546
OY 1267 TCCTCTCATTATGAGCCTTCAACCAACCAAAAGTGGCCCCAACCAGAGGTGTTG 1326
DB 1547 TCCTCTCATTATGAGCCTTCAACCAACCAAAAGTGGCCCCAACCAGAGGTGTTG 1606
OY 1327 ACCCTTCCGTTTTCACACCGGTTCTGCTCAACACAGCAGCTTCTGCTCTCAG 1386
DB 1607 ACCCTTCCGTTTTCACACCGGTTCTGCTCAACACAGCAGCTTCTGCTCTCAG 1666
OY 1387 GAGGATCAAGAACTGCATTTGGGAAACAATTTGCCATGAACGAGCTGAAGTGGCCACG 1446
DB 1667 GAGGATCAAGAACTGCATTTGGGAAACAATTTGCCATGAACGAGCTGAAGTGGCCACG 1726
OY 1447 CCCTGACCCCTGCTCGCTTTGAGCTGCTGCTGATCCCAACAGGATCCCAATTG 1506
DB 1727 CCCTGACCCCTGCTCGCTTTGAGCTGCTGCTGATCCCAACAGGATCCCAATTG 1786
OY 1507 CACGACTGTGTGTAATAATCCAAATGGAATCCACTGCGTCTCAGAGGCTCCCTAAC 1566
DB 1787 CACGACTGTGTGTAATAATCCAAATGGAATCCACTGCGTCTCAGAGGCTCCCTAAC 1846
OY 1567 CTTGTGAAGACAAGACCAAGCTTTGA 1592
DB 1847 CTTGTGAAGACAAGACCAAGCTTTGA 1872
```

## RESULT 9

ABK50068  
ID ABK50068 standard; DNA; 2116 BP.

XX AC ABK50068;

XX DT 15-JUL-2002 (first entry)

XX DB DNA encoding mouse Cyp 4A12 protein.

XX KW Mouse; Cyp 4A14; hypertension; Cyp 4A11; testosterone inhibitor; obesity;

KW lipid metabolism disease; pancreatic dysfunction; type II diabetes;

KW cardiovascular disease; Cyp 4A12; gene; ds.

XX OS Chimeric - Mus sp.

OS Chimeric - Synthetic.

XX FH Key Location/Qualifiers

FT CDS 282..1808

FT /tag= a

FT /product= "Mouse Cyp 4A12 protein"

FT /transl\_except= (pos:1044..1046, aa:Gln)

XX PN WO200217856-A2.

XX PD 07-MAR-2002.

XX PF 29-AUG-2001; 2001WO-US26914.

XX PR 29-AUG-2000; 2000US-228947P.

XX PA (UYVA-) UNIV VANDERBILT.

XX PI Capdevila J, Waterman M, Holla V;

DR WPI; 2002-382929/41.  
DR P-PSDB; AAU79993.

XX Treating hypertension in an individual by inhibiting testosterone  
PT activity, enhancing Cyp 4A14 activity, or by inhibiting Cyp 4A11  
PT activity in the individual -

XX Claim 5; Page 69-70; 80pp; English.

XX The present invention relates to a new method of treating hypertension in  
CC an individual. The method comprises inhibiting testosterone activity,  
CC enhancing Cyp 4A14 activity, or inhibiting Cyp 4A11 activity in the  
CC individual. Enhancing Cyp 4A14 activity leads to inhibition of  
CC testosterone activity which in turn leads to inhibition of Cyp 4A11  
CC activity. The method of the invention can be used for treating  
CC hypertension in an individual. The molecules of the invention are also  
CC useful for treating disease states associated with lipid metabolism,  
CC pancreatic dysfunction, obesity, type II diabetes, and other  
CC cardiovascular diseases. The present nucleic acid sequence encodes the  
CC mouse Cyp 4A12 protein of the invention, as described above.

XX Sequence 2116 BP; 517 A; 568 C; 441 G; 590 T; 0 other;

Query Match 42.0%; Score 1082.8; DB 24; Length 2116;  
Best Local Similarity 76.9%; Pred. No. 5.4e-239;  
Matches 1375; Conservative 0; Mismatches 402; Indels 11; Gaps 4;

QY 11 GAGATCCAGCAGGTGCTGCACCATGAGTGTCTGTGCTGAGCCCGCAGCAGACTCTGGG 70  
DB 260 GTGATCCAGAGCTGTGTATCATGAGTGCCCTGTCTGAGCTCCATCAGATTCCAGG 319  
QY 71 TGATGTCTTGAATCCTCCAGCGGCTCCCTGCTCATTTGCTTGTGCTGATCAA 130  
DB 320 AAGCATCTTGAGTACCTTCAAGTAGCCTGTGTCTGAGCCTGCTCTGCTGCTTCAA 379  
QY 131 GCGAGTTCAGCTCTACCTGCACAGCAGTGTGCTCAAGCCCTCCAGCAGTCCCGTG 190  
DB 380 GACAGCCCAAGCTCTACCTGCACAGCAGTGTGCTCAAGTACTCAGCAGTCCCATC 439  
QY 191 CCCTCCCTCCCACTGCTGCTGGGCAATCCAGAGCTCCAAAGAGCCAGAGCTACA 250  
DB 440 CCCACCTTCTCACTGCTCTTTGGACACA-----AGATCTTAAAGACCAGACTTCA 493  
QY 251 ACGGATTCAGAAATGGGTGAGACATTCCTCAAGTGTCTCTCATTTGGCTATGGGAGG 310  
DB 494 AGATATCTTAAGTATGAGATTAAGATTTCCCAAGTGTCTTCCACAGTGTCTGGGAGG 553  
QY 311 CAAGTTCTGTTCAGCTCTATGACCTGACTATATGAGTGTATTTCTGGGAGATCAGA 370  
DB 554 CAAGTGGCGCATTCAGTGTATGACCTGACTACATGAGTGTATTTCTGGGAGATCAGA 613  
QY 371 CCGGAATCCCATGTTCTCTACAGATTCCTGCTCCATGATGGGTAGCGTGTCTCT 430  
DB 614 CCGAAAAGCTAATGTTCTCTACAGATTTCTAGCTCCCTGATGGGCGTGTCTTAT 673  
QY 431 GTTGAATGGGCAAGATGTTCCAGCATGACGGATGCTGACCCCAAGCTTCCACTATGA 490  
DB 674 GCTGATGGAACAGACATGTTTACGACCCGAAATGTTGACCCCACTTCCACTATGA 733  
QY 491 CATCTGAAGCCCTATGTGGGGCTCATGGCAGACTCTGTAGAGATGCTGACAAATG 550  
DB 734 CATCTGAAGCCCTTATACGGAATCATGGCAGACTCTGTTCGTATGCTGATTAATG 793  
QY 551 GGAAGCTCTCTGGCCAGATTCCTCTGAGGTCTTTCAAGCAGTCTCTGATGAC 610  
DB 794 GGAACAGATTTGGCCAGATTCACCTGAGATCTTTCAGACATCACTTGAATGAC 853  
QY 611 CTTGACACCATCATGAGTGTCTTCAAGCATCAGGCGACATCCAGTGAAGAGAA 670  
DB 854 CTTGACACCATCATGAGTGTCTTCAAGCATCAGGCGAGTGTCCAGTTGAAGAGAA 913  
QY 671 TTCTCACTCTATACAGGCGCATTAAGTGAAGTGAACCACTGTTTTCCTGAG 730

DB 914 ATACAGTCTATATCCAGCAGTGAAGACCTGAACGATCTCGTTTTTCCCGTGTGCG 973  
QY 731 GAATGCTTTTACCAGAAATGACACCATCTACAGCTGACCTGTGCGCCGCTGACACA 790  
DB 974 GAACATCTTACCTGAATGACATCATCTACAGAGTGTCTCTAATGCTGCAAGCTTA 1033  
QY 791 CCGGCTGCTCCAGCTGGCCATCAGACACAGACCAAGTATCCAACTGAGAAAGCTCA 850  
DB 1034 CAGTGCCTGCMAACTTGCATGATATCACACAGACCAAGTATCAATCAAGAGATTC 1093  
QY 851 ACTACAGAGAGGGGAGCTGAGAAATCAAGAGAAAGAGCATTGATTTCTGGA 910  
DB 1094 ACTTCAGATGAGAAAGTTGAAAGCTTAAGAAAGAAAGCGATTGATTTCTGGA 1153  
QY 911 TATCTCTCTTGGCCAAATGAGAAATGGAGCACTTGTCAAGCAAGACCTCGTGC 970  
DB 1154 CATCTCTATTTGGCCAGATGGAATGAAATGAAAGCTTATCTGATTAAGACCTTGTGC 1213  
QY 971 TGAAGTGAACAGCTTCACTTTGAGGGCCACGACACAGCCAGTGGATCTCTGAT 1030  
DB 1214 TGAAGTGAATCACTTCACTTTGAGGGCCATGACACCAAGTATCTCTGAT 1273  
QY 1031 CCTCTATGCTCTGCGCCACACACCCCAAGCATCAGAGAGTGCCTGGAGATCCACAG 1090  
DB 1274 CTCTATGCTTTGGCCCAAAATCTGAAATCAAGAGATGAGAAAGATCCAAAG 1333  
QY 1091 CCTCTGGGTGATGAGCTTCCATCACTGGAACCACTGACCAAGATGCCCTTACACAC 1150  
DB 1334 TCTCTAGAGATGAGCTTCTATCACTGGAATGACCTGACAAAGATGCCCTATATAC 1393  
QY 1151 CATGTGATTAAGAGGCACTGAGGCTTACCCACCGGTGCCAGGATTTGACAGAGCT 1210  
DB 1394 CATGTGATTAAGAGGCGCTGAGGATCTACCTCTGTACCAAGTGTGAGCAGAGACT 1453  
QY 1211 CAGCATCCCGTACCTTCCCTGATGGGCGCTCTTCCCAAGATCATGCTCTCT 1270  
DB 1454 CAGCTGACCTGTACCTTCCAGATGAGCGTCTTACCAAGATATCCATGATATGCT 1513  
QY 1271 CTGCAATTAATGCGCTTACCAACCAACCAAGTGTGGCCCAACCAAGGTTTGAACC 1330  
DB 1514 CTCTTTATGCGCTTCAATCAACCAACCAAGTGTGGCCCAATCCAGAGTGTGATCC 1573  
QY 1331 TTTCGTTTTGCAACCGGCTTCTGCTCAACACAGCCAGCTTCTGCTCTCAGAGG 1390  
DB 1574 TTCTGATTTGCAACAGGCTTCTCCGACAGCCAGCTCATTTCTGCTCTCAGAGG 1633  
QY 1391 ATCAAGAACTGATTTGGAACAATTTGCCATGAACGAGTGAAGTGGCCACGGCT 1450  
DB 1634 AGCAAGAACTGATTTGGAACAATTTGCCATGAACGAGTGAAGTGGCTGAGGCT 1693  
QY 1451 GACCTGCTCCGCTTGAAGTGTGCTGCTGATCCCAACAGATCCCATTCAGAG 1510  
DB 1694 GACCTGCTCCGCTTGAAGTGTGCTGCTGATCCCAACAGATCCCATTCAGAG 1753  
QY 1511 ACTTGTGTAATCCAAATGGAATCCAGCTGCTGCTGAGAGGCTCCCTAACCTTG 1570  
DB 1754 AATTGTGTAATCCAAATGGAATCCAGCTGCTGCTGAGAGGCTCCCTAACCTTG 1813  
QY 1571 TGAAGCAAGACAGCTTTGAGGCTCCACCTGCGCTGCTGCTGAGAGGCTCCCT 1630  
DB 1814 ACAGAGCAAG--ACAGCTAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1871  
QY 1631 TCTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1689  
DB 1872 CTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1931  
QY 1690 CACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1747  
DB 1932 CACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1991  
QY 1748 CCCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1795  
DB 1992 CCCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2039





[illegible]

XX	RESULT 11
XX	ABK50071
ID	ABK50071 standard; DNA; 21990 BP.
XX	
AC	ABK50071;
XX	
DT	15-JUL-2002 (first entry)
XX	
DE	Human CYP 4A11 genomic DNA sequence.
XX	
KM	Human; CYP 4A14; hypertension; CYP 4A11; testosterone inhibitor; obesity;
KM	lipid metabolism disease; pancreatic dysfunction; type II diabetes;
XX	cardiovascular disease; ds.
XX	
OS	Chimeric - Homo sapiens.
OS	Chimeric - Synthetic.
XX	
PN	WO200217856-A2.
XX	
PD	07-MAR-2002.
XX	
PF	29-AUG-2001; 2001WO-US26914.
XX	
PR	29-AUG-2000; 2000US-228947P.
XX	
PA	(UYVA-) UNIV VANDERBILT.
XX	
PI	Capdevila J, Waterman M, Holla V;
XX	
DR	WPI; 2002-382929/41.
XX	

PT	Treating hypertension in an individual by inhibiting testosterone
PT	activity, enhancing Cyp 4A14 activity, or by inhibiting Cyp 4A11
PT	activity in the individual -
XX	
PS	Claim 11; Page 74-80; 80pp; English.
XX	
CC	The present invention relates to a new method of treating hypertension in
CC	an individual. The method comprises inhibiting testosterone activity,
CC	enhancing Cyp 4A14 activity, or inhibiting Cyp 4A11 activity in the
CC	individual. Enhancing Cyp 4A14 activity leads to inhibition of
CC	testosterone activity which in turn leads to inhibition of Cyp 4A11
CC	activity. The method of the invention can be used for treating
CC	hypertension in an individual. The molecules of the invention are also
CC	useful for treating disease states associated with lipid metabolism,
CC	pancreatic dysfunction, obesity, type II diabetes, and other
CC	cardiovascular diseases. The present nucleic acid sequence represents
CC	the human Cyp 4A11 genomic DNA sequence of the invention, as described
CC	above.
XX	
XX	
Sequence	21990 BP; 1784 A; 2031 C; 1682 G; 1899 T; 14594 other;

Query Match	37.6%;	Score 968.4;	DB 24;	Length 21990;
Best Local Similarity	98.0%;	Pred. No. 2.6e-212;		
Matches 1002; Conservative	0;	Mismatches 16;	Indels 4;	Gaps 2;

QY	1381	TCTCAGAGGATCAAGGAATCTGATTTGGGAAACAATTGGCATGAACGAGCTGAAGGTGG	1440
Db	13759	TCTCCACTGGCCCCAGGAATCGATCGGGAAACAATTGGCCATGAACGAGCTGAAGGTGG	13818
QY	1441	CCACGGCCCTGACCCCTGCTCCGCTTTGAGCTGCTGCCTGATCCACCAGATCCCATCC	1500
Db	13819	CCACGGCCCTGACCCCTGCTCCGCTTTGAGCTGCTGCCTGATCCACCAGATCCCATCC	13878
QY	1501	CCATTGCAAGACTTGTGTTGAATCCAAAAATGGAATCACTGCGTCTCAGGAGGCTCC	1560
Db	13879	CCATTGCAAGACTTGTGTTGAATCCAAAAATGGAATCACTGCGTCTCAGGAGGCTCC	13938
QY	1561	CTAACCCCTGTGAAGACAAGACCAAGCTTTGAGGAGCTCCACTGCGTCTGTCTTCT	1620
Db	13939	CTAACCCCTGTGAAGACAAGACCAAGCTTTGAGGAGCTCCACTGCGTCTGTCTTCT	13998
QY	1621	GACCCCGGCTTCTGTCTCCCTTCTCTGTCTGCCATATCTGTCTTCTGTCTGCCACTTC	1680
Db	13999	GACCCCGGCTTCTGTCTCCCTTCTCTGTCTGCCATATCTGTCTTCTGTCTGCCACTTC	14058
QY	1681	CCTTCTTCCCACTGCGCTGTGCCCCAGTCTGCGCTGCGCTTCTCTCTCACCTTCT	1740
Db	14059	CCTTCTTCCCACTGCGCTGTGCCCCAGTCTGCGCTGCGCTTCTCTCTCACCTTCT	14118
QY	1741	CCAGGCTCCCTAACCTGCTTGTCTTACTGTCTCTTACCACCTGTATCTTGTGGGAGA	1800
Db	14119	CCAGGCTCCCTAACCTGCTTGTCTTACTGTCTCTTACCACCTGTATCTTGTGGGAGA	14178
QY	1801	AAAGCTGAGTGTGGGAGAGCTGAGGCCGAGCTTGCACTGTGACATTAATGTAAAGAG	1860
Db	14179	AAAGCTGAGTGTGGGAGAGCTGAGGCCGAGCTTGCACTGTGACATTAATGTAAAGAG	14238
QY	1861	TCTTGAATCATGTCCAGGATCCAGGGTCTAAAAACCCCTGTGGCCTTTGGAACACCAAGC	1920
Db	14239	TCTTGAATCATGTCCAGGATCCAGGGTCTAAAAACCCCTGTGGCCTTTGGAACACCAAGC	14298
QY	1921	TCTGTGCTGAAGGGGTGAAGGCTTACCTTGAACGACCATTAATCTAAGCCCCGGGGCATAAA	1980
Db	14299	TCTGTGCTGAAGGGGTGAAGGCTTACCTTGAACGACCATTAATCTAAGCCCCGGGGCATAAA	14358
QY	1981	CCCCTGATGCTTGATAGATCCAGGCTCGTGCTCTGGAATGTGTCTGACTTGCTG	2040
Db	14359	CCCCTGATGCTTGATAGATCCAGGCTCGTGCTCTGGAATGTGTCTGACTTGCTG	14417
QY	2041	GCTCCTGTGCTTGTCTCTCCAGATCAATGTATCTTGAGTTAAAGAACCTGCTCTC	2100
Db	14418	CCTCCTGTGCTTGTCTCTCCAGATCAATGTATCTTGAGTTAAAGAACCTGCTCTC	14477

QY 2101 CATATCTCAAGTACAGACAGATGCTAAACCGTCACAGCTGTAAATTGTGTCTTAAT 2160  
DB 14478 CATATCTCAAGTACAGACAGATGCTAAACCGTCACAGCTGTAAATCATGTGTTAAT 14537  
QY 2161 GCAACATGCCCTTTGACCCACCCCATTTCTACCACTGTTTCTTTGTTGATCAACA 2220  
DB 14538 GCAACATGCCCTTTGAC---CCCCATTTCTACCACTGTTTCTTTGTTGATCAACA 14594  
QY 2221 ATAAATAATCTGCACTTCCAGAGCTGGGGGCTTCAAGCTTCCATCTTAGCTTTGGCG 2280  
DB 14595 ATAAATAATCTGCACTTCCAGAGCTGGGGGCTTCAAGCTTCCATCTTAGCTTTGGCG 14654  
QY 2281 CCTGGAACCCACTTTCTCTCTCAAACTGTTCTTTCTCACTGCTTGAATCTGCCGACTT 2340  
DB 14655 CCTGGAACCCACTTTCTCTCTCAAACTGTTCTTTCTCACTGCTTGAATCTGCCGACTT 14714  
QY 2341 TGTCAACCCCAAGACCTGTTGGGTCTGAACACCCCAATCCCTGAATCTCAACCCA 2400  
DB 14715 TGTCAACCCCAAGACCTGTTGGGTCTGAACACCCCAATCCCTGAATCTCAACCCA 14774  
QY 2401 CC 2402  
DB 14775 CC 14776

RESULT 12  
ABK50067  
ID ABK50067 standard; DNA; 4123 BP.

XX ABK50067;

DT 15-JUL-2002 (first entry)

DE DNA encoding mouse Cyp 4A14 protein.

XX Mouse; Cyp 4A14; hypertension; Cyp 4A11; testosterone inhibitor; obesity;

KW lipid metabolism disease; pancreatic dysfunction; type II diabetes;

KM cardiovascular disease; gene; ds.

XX Chimeric - Mus sp.

OS Chimeric - Synthetic.

XX Key

FT CDS

FT 1637..3161

FT /tag= a

FT /product= "Mouse Cyp 4A14 protein"

FT /transl\_except= (pos:1649..1651, aa:Val)

FT /transl\_except= (pos:1832..1835, aa:Asp)

FT /note= "This codon has an apparent 1 nucleotide

FT insertion which alters the reading frame"

XX WO200217856-A2.

XX PD 07-MAR-2002.

XX PF 29-AUG-2001; 2001WO-US26914.

XX PR 29-AUG-2000; 2000US-228947P.

XX PA (UYVA-) UNIV VANDERBILT.

XX PI Capdevila J, Waterman M, Holla V;

XX DR WPI; 2002-382929/41.

XX P-PSDB; AAU79992.

XX PT Treating hypertension in an individual by inhibiting testosterone

XX PT activity, enhancing Cyp 4A14 activity, or by inhibiting Cyp 4A11

XX PT activity in the individual -

XX PS Claim 1; Page 65-66; 80pp; English.

XX CC The present invention relates to a new method of treating hypertension in

CC an individual. The method comprises inhibiting testosterone activity,  
CC enhancing Cyp 4A14 activity, or inhibiting Cyp 4A11 activity in the  
CC individual. Enhancing Cyp 4A14 activity leads to inhibition of  
CC testosterone activity which in turn leads to inhibition of Cyp 4A11  
CC activity. The method of the invention can be used for treating  
CC hypertension in an individual. The molecules of the invention are also  
CC useful for treating disease states associated with lipid metabolism,  
CC pancreatic dysfunction, obesity, type II diabetes, and other  
CC cardiovascular diseases. The present nucleic acid sequence encodes the  
CC mouse Cyp 4A14 protein of the invention, as described above.

SO Sequence 4123 BP; 1130 A; 860 C; 867 G; 1265 T; 1 other;  
Query Match 37.5%; Score 966; DB 24; Length 4123;  
Best Local Similarity 73.5%; Pred. No. 5.1e-212;  
Matches 1298; Conservative 0; Mismatches 440; Indels 28; Gaps 4;

QY 30 ACCATGATGTCTCTGTGCTGAGCCCAAGACAGACTCCGTGGTGTCTCTGGAATCCTC 89  
DB 1634 ACCATGATGTCTCTGTGCTGAGCCCAAGACAGACTCCGTGGTGTCTCTGGAATCCTC 1693  
QY 90 CAAGCGCTCCCTGCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 149  
DB 1694 CAATGGGCTTCTTCTGCTCAGCTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1753

QY 150 CACAGGATGTGCTGCTCAAAAGCCCTCCAGCAGTTCCTGCTGCTGCTGCTGCTGCTGCT 209  
DB 1754 CCAAGGATGTGCTGCTCAAAAGCCCTCCAGCAGTTCCTGCTGCTGCTGCTGCTGCTGCT 1813

QY 210 TTGGGATCATCCAGAGCTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 269  
DB 1814 TTGGGATCATCCAGAGCTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1865

QY 270 GAGACATTTCCCAAGTCCCTGCTCATTTGGCTATGGGAGGCAAGTTGCTGCTGCTGCTG 329  
DB 1866 GAGACATTTCCCAAGTCCCTGCTCATTTGGCTATGGGAGGCAAGTTGCTGCTGCTGCTG 1925

QY 330 TATGACCTGACTATATGAAGTATTTCTGGGAGATCAACCCGAATCCCATGCTTCC 389  
DB 1926 TATGACCTGACTATATGAAGTATTTCTGGGAGATCAACCCGAATCCCATGCTTCC 1985

QY 390 TACAGATTCCTGCTCCATGATGGGTACGGCTTCTCTGTTGAATGGGAGAGACATGG 449  
DB 1986 TACAGATTCCTGCTCCATGATGGGTACGGCTTCTCTGTTGAATGGGAGAGACATGG 2045

QY 450 TTCCAGATTCAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 509  
DB 2046 TTCCAGATTCAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2105

QY 510 GGGCTGATGAGCAACTCTGTAGAGTGTGATGATGATGATGATGATGATGATGATGATG 569  
DB 2106 GGGCTGATGAGCAACTCTGTAGAGTGTGATGATGATGATGATGATGATGATGATGATG 2165

QY 570 GATTCCTCTTGAAGTCTTTCAACAGCTCTCTGATGATGATGATGATGATGATGATGATG 629  
DB 2166 GATTCCTCTTGAAGTCTTTCAACAGCTCTCTGATGATGATGATGATGATGATGATGATG 2225

QY 630 TGTGCTTCAAGCATCAGGAGAGATCCAGGTGAGACAGAAATTTCTAGTCTTACATACAG 689  
DB 2226 TGTGCTTCAAGCATCAGGAGAGATCCAGGTGAGACAGAAATTTCTAGTCTTACATACAG 2285

QY 690 GCCATTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 749  
DB 2286 GCCATTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2345

QY 750 GACACATCTACAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 809  
DB 2346 GACACATCTACAGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2405

QY 810 CATCAGACACAGACCAAGTATCCAACTGAGGAGAGGCTCAACTACAGAAAGAGGGGAG 869  
DB 2406 CATCAGACACAGACCAAGTATCCAACTGAGGAGAGGCTCAACTACAGAAAGAGGGGAG 2465







PR 17-SEP-1998; 98WO-US19437.  
PR 07-OCT-1998; 98WO-US21141.  
PR 29-OCT-1998; 98WO-US22991.  
PR 29-OCT-1998; 98WO-US22992.  
PR 20-NOV-1998; 98WO-US24855.  
PR 01-DEC-1998; 98WO-US25108.  
PR 05-JAN-1999; 99WO-US00106.  
PR 08-MAR-1999; 99WO-US05028.  
PR 10-MAR-1999; 99WO-US05190.  
PR 20-APR-1999; 99WO-US0615.  
PR 14-MAY-1999; 99WO-US10733.  
PR 02-JUN-1999; 99WO-US12252.  
PR 01-SEP-1999; 99WO-US20111.  
PR 08-SEP-1999; 99WO-US20594.  
PR 13-SEP-1999; 99WO-US20944.  
PR 15-SEP-1999; 99WO-US21090.  
PR 15-SEP-1999; 99WO-US21547.  
PR 05-OCT-1999; 99WO-US23089.  
PR 29-NOV-1999; 99WO-US28214.  
PR 30-NOV-1999; 99WO-US28313.  
PR 30-NOV-1999; 99WO-US28409.  
PR 01-DEC-1999; 99WO-US28301.  
PR 01-DEC-1999; 99WO-US28634.  
PR 02-DEC-1999; 99WO-US28551.  
PR 02-DEC-1999; 99WO-US28564.  
PR 02-DEC-1999; 99WO-US28565.  
PR 16-DEC-1999; 99WO-US30095.  
PR 20-DEC-1999; 99WO-US30911.  
PR 22-DEC-1999; 99WO-US30999.  
PR 22-DEC-1999; 99WO-US30720.  
PR 30-DEC-1999; 99WO-US31243.  
PR 30-DEC-1999; 99WO-US31274.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00277.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 18-FEB-2000; 2000WO-US04342.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 24-FEB-2000; 2000WO-US05004.  
PR 01-MAR-2000; 2000WO-US05601.  
PR 02-MAR-2000; 2000WO-US05746.  
PR 02-MAR-2000; 2000WO-US05841.  
PR 10-MAR-2000; 2000WO-US06319.  
PR 15-MAR-2000; 2000WO-US06884.  
PR 20-MAR-2000; 2000WO-US07377.  
PR 21-MAR-2000; 2000WO-US07532.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 11-AUG-2000; 2000WO-US22031.  
PR 23-AUG-2000; 2000WO-US23522.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 10-NOV-2000; 2000WO-US30873.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 20-DEC-2000; 2000WO-US34956.  
PR 28-FEB-2001; 2001WO-US06520.  
PR 01-MAR-2001; 2001WO-US06666.  
PR 25-MAY-2001; 2001WO-US17092.  
PR 01-JUN-2001; 2001WO-US17800.  
PR 20-JUN-2001; 2001WO-US19692.  
PR 22-JUN-2001; 2001WO-US20116.  
PR 29-JUN-2001; 2001WO-US21066.  
PR 09-JUL-2001; 2001WO-US21735.  
PR 20-DEC-2000; 2000US-0747259.  
PR 28-FEB-2001; 2001US-0796498.  
PR 09-MAR-2001; 2001US-0802706.  
PR 14-MAR-2001; 2001US-0808689.

PR 22-MAR-2001; 2001US-0816744.  
PR 05-APR-2001; 2001US-0828366.  
PR 10-MAY-2001; 2001US-0854208.  
PR 10-MAY-2001; 2001US-0854280.  
PR 18-MAY-2001; 2001US-0860216.  
PR 25-MAY-2001; 2001US-0866028.  
PR 25-MAY-2001; 2001US-0866034.  
PR 01-JUN-2001; 2001US-0872035.  
PR 05-JUN-2001; 2001US-0874503.  
PR 14-JUN-2001; 2001US-0882636.  
PR 19-JUN-2001; 2001US-0886342.  
PR 21-JUN-2001; 2001US-0887879.  
PR 18-JUL-2001; 2001US-0908827.  
PR 06-AUG-2001; 2001US-0924419.  
PR 09-AUG-2001; 2001US-0927796.  
PR 16-AUG-2001; 2001US-0931836.  
PR 19-DEC-2001; 2001US-0028072.  
XX  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX  
XX WPI: 2003-332040/31.  
DR P-PSDB; ABU66623.  
XX  
XX  
PT New secreted and transmembrane PRO nucleic acids, useful for gene  
PT therapy, in chromosome and gene mapping, as chromosome markers, in  
PT tissue typing, and in chromosome identification  
XX  
XX  
PS Claim 2; Fig 107; 660pp; English.  
XX  
XX The present invention relates to the isolation of novel human PRO  
CC polypeptides, and the polynucleotide sequences encoding them. The  
CC PRO polypeptides are secreted and transmembrane proteins. The PRO  
CC polypeptides are useful for detecting other PRO polypeptides, for  
CC linking bioactive molecules to cells expressing PRO polypeptides,  
CC for modulating biological activities of cells expressing PRO  
CC polypeptides, and for for identifying agonists or antagonists.  
CC The PRO polypeptides are useful for for stimulating the release of  
CC tumour necrosis factor (TNF)-alpha from human blood, for stimulating  
CC the proliferation or differentiation of chondrocytes, and detecting the  
CC presence of tumours. The polynucleotide sequences encoding PRO  
CC polypeptides are useful as hybridisation probes, in chromosome and  
CC gene mapping, in the generation of antisense RNA and DNA, in the  
CC preparation of PRO polypeptides, for generating transgenic animals or  
CC knockout animals, for the genetic analysis of individuals with genetic  
CC disorders, and in gene therapy. ACA03603-ACA03877 represent cDNAs  
CC encoding the human PRO polypeptides of the invention.  
CC Note: The sequence data for this patent was obtained in electronic  
CC format directly from the USPTO web site at  
CC [seqdata.uspto.gov/psipsdIDBntry.html](http://seqdata.uspto.gov/psipsdIDBntry.html).  
XX  
SQ Sequence 2343 BP; 607 A; 581 C; 525 G; 630 T; 0 other;  
  
Query Match 20.7%; Score 534; DB 25; Length 2343;  
Best Local Similarity 61.5%; Pred. No. 1e-112;  
Matches 912; Conservative 0; Mismatches 560; Indels 12; Gaps 3;  
  
QY 95 GGCCTGCTGCTCATTTCTTCTGCTGCTGATCAAGCAGTTCAAGCTTACCTGACAG 154  
DB 281 GGGCTGCTGTTCTGCTGCTGCGCCCTGGGGCTGCTGCAGGCCATTAACTGTACCTGCGAG 340  
  
QY 155 GCAGTGGCTGCTCAAGCCCTCCAGCAGTTCCCGTGGCCCTCCCTCCCACTGGCTTTCCG 214  
DB 341 GCAGCGGCTGCTGCGGAGCTGCGCCCTTCCAGCGGCCCCACCCACTGGTTCTTGG 400  
  
QY 215 GCACATCCAGAGCTCAACAGAGCCAGAGCTACACGGAATTCAGAAATGGGTGAGAC 274  
DB 401 GCA--CCAGAGTTTATTCAGATGATACATGAGAAAGCTTGAGAAATTAATGAAAA 457  
  
QY 275 ATTCCCAAGTGCCTGCTCTCATTTGGCTATGCGGAGGCAAGTTCTGTCCAGCTCTATGA 334

```
Db 458 ATACCTCGTCCCTTCCCTTCTTGAGTTGGCCCTTTCAGGCATTCTTCTGATCTATGA 517
Qy 335 CCCTGACTATATGAGGTGATTCTGGGGAGATCAGACCCGAAATCCATGGTCTCTACAG 394
Db 518 CCCAGACTATGCAAGACACTTCTGAGCAGAACAGATCCAGTCCAGTAACCTGCAGAA 577
Qy 395 ATTCCTGCTCCATGATGGGTAGCGCTGCTCTGTTGAATGGGAGACATGGTTCCA 454
Db 578 ATTCTCACTCCACTTCTTGAAAAAGACTAGCGGCTCTAGACGACCCAGTGTTC 637
Qy 455 GCATCGACGAGTGTGATACCCCGACCTTCCAATGACATCTGAGCCCTATGTGGGCT 514
Db 638 GCATCGTCCCTACTAATCTCTGATTCATTTTAACATCTGAAAGCATACATTGAGGT 697
Qy 515 CATGCGAGACTCTGTACGAGTGTGTCGACAAATGGGAAGACTCTTGGC--CAGGA 571
Db 698 GATGCTCATCTCTGAAATGATGCTGATTAAGTGGAGAGATTTCAGACACTCAGGA 757
Qy 572 TTCCCTCTGAGGCTCTTTCAGCAGCTCTCTGATGACCTGACACCATCATGAAGTG 631
Db 758 CACAAGCTGGAGGTCTATGAGCACAATCTGATGTCTGATATATCATGAAGATG 817
Qy 632 TGCTTTAGCCATCAGGGCAGCATCCAGGTGGAACAGGAATTCAGTCTTACATACAGGC 691
Db 818 CGCTTTAGCAGAGAGACCAACTGCGCAGACAAACAGACCCCATGATCTTATGCAAAAGC 877
Qy 692 CATTAAGTACCTGAAACAACCTGGTTTTCCTGCTGAGGAATGCTTTACCAAGAAATGA 751
Db 878 CATATTTGAATCAGCAAAATCATATTTTCAACCGCTGTACAGTTGTTGTATCACAGTGA 937
Qy 752 CACCATCTACAGCTGACCTCTGCTGGCGCTGGAACACACCGCGCTGCGAGCTGGCCCA 811
Db 938 CATATTTTCAACTCAGCCCTCAGGGCTACCGCTTCAGAACTTAAGCCGAGTGTGA 997
Qy 812 TCAGCACAGACCAAGTGATCCAATGAGAGAGGCTCAACTACAGAAAGAGGGGAGCT 871
Db 998 TCAGTACACAGATACATAATCCAGAAAGAAAGAAATCCCTCCAGGCTGGGGTAAAGCA 1057
Qy 872 GGAGAAATCAAGAGAGAGGCAATTTTGATTTCTGATATCTCTCTTGCCCAAAAT 931
Db 1058 GGATTAACACTCCGAAGAGAGGATCCAGGATTTCTGATATTTCTTCTGCCAAGGA 1117
Qy 932 GGAGAAATGGAGCATCTTGTACAGACAGACCTCGGTGCTGAGTGAACAGTTCAATGTT 991
Db 1118 TGAAGTGTAGCAGCTTCTCAGATATGATGTACACTCTGAAGTGAACACATTCTGTT 1177
Qy 992 TGAGGGCCACGACACACAGCCAGTGGATCTCTGATCTCTATGCTTGCCACACA 1051
Db 1178 GGCAGGACATGACACTTGGCAGACAGATCTCTGATCTCTTACTGCTGCTGAA 1237
Qy 1052 CCCCAAGCATCAGAGAGGTGCGGGAGAGATCCACAGCTCCTGGGTGATGAGCCTC 1111
Db 1238 CCTGAGCATCAAGAGAGATGCCGGAGAGAGGTCAAGGGCATCTGGGGATGGGTCTTC 1297
Qy 1112 CATCACTGTGAACCACTTGAACCAAGTGCCTTACACCAACCATGTGCAATTAAGAGGCACT 1171
Db 1298 TATCACTTGGGACCACTGGGTGAGATGTGTACACCACAATGTGCAATCAAGGAGACGTG 1357
Qy 1172 GAGGCTTAACCAACCGGTGCGAGCATTGGCAGAGAGCTCAGCACTCCGTCACTTCCC 1231
Db 1358 CCGATTGATCTTGCAGTCCCGTCCATTTCCAGAGATCTCAGCAAGCCACTTACCTCCC 1417
Qy 1232 TGATGGCGCTCTTGCCCAAGGTATCATGTCTCTCTCTCAATTAAGCCTTACCA 1291
Db 1418 AGATGATGACATGTGCTGACAGGATCACCGTGTCTTATGATTTGGGCTTTCACCA 1477
Qy 1292 CAACCCAAAGTGTGGCCCAACCCAGAGGTGTTGACCTTCCGTTTTCACCGGGTTC 1351
Db 1478 CAACCTGTCTGTGAAAAAACCAAGGTCTTTGACCCCTTGAGGTCTCTCAGGAGAA 1537
Qy 1352 TGCTCAACACAGCCAC-----GCTTCTGCGCTTCTCAGGAGATCAAGGAATGCAAT 1405
```

```
Db 1538 TTCTGATCAGAGACACCCCTATGCTACTTACATTTCTCAGCTGATCAAGAACTGCAT 1597
Qy 1406 TGGGAACAATTTGCCATGAACGAGCTGAAGGTGGCCACGGCCCTGACCTGCTCCGCTT 1465
Db 1598 TGGGACAGAGTTTGCCATGATTTGAGTTAAAGGTAACCATTTGCCCTGATTTCTCCACTT 1657
Qy 1466 TGAGCTGCTGCTGATCCCAACAGGATCCCAATCCCATTTGACAGCACTGTGTTGAATC 1525
Db 1658 CAGAGTACTCCAGACCCCAACAGGCTCTTACTTTCCCAACCATTTTATCTCAAGCC 1717
Qy 1526 CAAAAATGAATCCACTGCTCTCAGAGAGGCTCCCTAACCTT 1569
Db 1718 CAAGATGGATGTATTGTGACCTGAAGAACTCTCTGAATGTT 1761

RESULT 15
ACA04077
ID ACA04077 standard; cDNA, 2343 BP.
XX
AC ACA04077;
XX
DT 27-MAY-2003 (first entry)
XX
DE Human cDNA encoding a secreted/transmembrane protein, SEQ ID 107.
XX
KW Human; ss; gene; secreted protein; transmembrane protein; PRO;
KW inflammatory disease; organ failure; atherosclerosis; cardiac injury;
KW infertility; birth defects; premature aging; AIDS; biosensor;
KW acquired immunodeficiency syndrome; cancer; diabetic complication;
KW bioreactor; tumour.
XX
OS Homo sapiens.
XX
PN US2003032155-A1.
XX
PD 13-FEB-2003.
XX
PF 03-MAY-2002; 2002US-0137865.
XX
PR 31-MAR-1997; 97WO-US05230.
PR 12-JUN-1998; 98WO-US12456.
PR 14-JUL-1998; 98WO-US14552.
PR 28-AUG-1998; 98WO-US17888.
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19093.
PR 14-SEP-1998; 98WO-US19094.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 07-OCT-1998; 98WO-US21141.
PR 29-OCT-1998; 98WO-US22991.
PR 29-OCT-1998; 98WO-US22992.
PR 20-NOV-1998; 98WO-US24855.
PR 01-DEC-1998; 98WO-US25108.
PR 05-JAN-1999; 99WO-US00106.
PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99WO-US05190.
PR 20-APR-1999; 99WO-US08615.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
```

PR	02-DEC-1999;	99WO-US28565.
PR	16-DEC-1999;	99WO-US30095.
PR	20-DEC-1999;	99WO-US30911.
PR	20-DEC-1999;	99WO-US30991.
PR	22-DEC-1999;	99WO-US30720.
PR	30-DEC-1999;	99WO-US31243.
PR	30-DEC-1999;	99WO-US31274.
PR	05-JAN-2000;	2000WO-US00219.
PR	06-JAN-2000;	2000WO-US00277.
PR	06-JAN-2000;	2000WO-US00376.
PR	11-FEB-2000;	2000WO-US03565.
PR	18-FEB-2000;	2000WO-US04341.
PR	18-FEB-2000;	2000WO-US04342.
PR	22-FEB-2000;	2000WO-US04414.
PR	24-FEB-2000;	2000WO-US04914.
PR	24-FEB-2000;	2000WO-US05004.
PR	01-MAR-2000;	2000WO-US05601.
PR	02-MAR-2000;	2000WO-US05746.
PR	02-MAR-2000;	2000WO-US05841.
PR	10-MAR-2000;	2000WO-US06319.
PR	15-MAR-2000;	2000WO-US06884.
PR	20-MAR-2000;	2000WO-US07377.
PR	21-MAR-2000;	2000WO-US07532.
PR	30-MAR-2000;	2000WO-US08439.
PR	17-MAY-2000;	2000WO-US13705.
PR	22-MAY-2000;	2000WO-US14042.
PR	30-MAY-2000;	2000WO-US14941.
PR	02-JUN-2000;	2000WO-US15264.
PR	28-JUL-2000;	2000WO-US20710.
PR	11-AUG-2000;	2000WO-US22031.
PR	23-AUG-2000;	2000WO-US23522.
PR	24-AUG-2000;	2000WO-US23328.
PR	08-NOV-2000;	2000WO-US30952.
PR	10-NOV-2000;	2000WO-US30873.
PR	01-DEC-2000;	2000WO-US34256.
PR	20-DEC-2000;	2000WO-US34956.
PR	28-FEB-2001;	2001WO-US06520.
PR	01-MAR-2001;	2001WO-US06666.
PR	25-MAY-2001;	2001WO-US17092.
PR	01-JUN-2001;	2001WO-US17800.
PR	20-JUN-2001;	2001WO-US19692.
PR	22-JUN-2001;	2001WO-US20116.
PR	29-JUN-2001;	2001WO-US21066.
PR	09-JUL-2001;	2001WO-US21735.
PR	20-DEC-2000;	2000US-0747259.
PR	28-FEB-2001;	2001US-0796498.
PR	09-MAR-2001;	2001US-0802706.
PR	14-MAR-2001;	2001US-0808689.
PR	22-MAR-2001;	2001US-0816744.
PR	05-APR-2001;	2001US-0828366.
PR	10-MAY-2001;	2001US-0854208.
PR	10-MAY-2001;	2001US-0854280.
PR	18-MAY-2001;	2001US-0860216.
PR	25-MAY-2001;	2001US-0866028.
PR	25-MAY-2001;	2001US-0866034.
PR	01-JUN-2001;	2001US-0872035.
PR	05-JUN-2001;	2001US-0874503.
PR	14-JUN-2001;	2001US-0882636.
PR	19-JUN-2001;	2001US-0886342.
PR	21-JUN-2001;	2001US-0887879.
PR	18-JUL-2001;	2001US-0908827.
PR	06-AUG-2001;	2001US-0924419.
PR	09-AUG-2001;	2001US-0927796.
PR	16-AUG-2001;	2001US-0931836.
PR	19-DEC-2001;	2001US-0028072.

DR P-PSDB; ABU66899.  
XX  
XX New secreted and transmembrane nucleic acids and polypeptides,  
PT designated as PRO, useful for treating inflammation, organ failure,  
PT atherosclerosis, cardiac injury, infertility, birth defects, premature  
PT aging, AIDS, or cancer  
XX  
XX  
PS Claim 2; Fig 107; 659pp; English.  
XX  
XX The invention relates to an isolated nucleic acid comprising, or which is  
CC at least 80% identical to, or the full-length coding sequence of, any of  
CC the 275 nucleotide sequences, encoding the corresponding PRO polypeptide  
CC (one of 275 secreted or transmembrane proteins). The nucleic acid  
CC further comprises the full-length coding sequence of the DNA deposited  
CC under American Type Culture Collection (ATCC) accession number in a list  
CC given in the specification. Also included are vectors and host  
CC cells for producing PRO proteins, PRO fusion proteins, anti-PRO  
CC antibodies, PRO extracellular domains and mature sequences, methods  
CC of detecting PRO proteins, methods for stimulating the release of  
CC TNF-alpha (tumour necrosis factor alpha) from human blood, cells  
CC (and the proliferation of differentiation of chondrocyte cells, the  
CC proliferation of, or gene expression in pericyte cells, the release or  
CC proteoglycans from cartilage, proliferation of inner ear utricular  
CC supporting cells, the proliferation of T-lymphocyte cells, the release  
CC of a cytokine from peripheral blood mononuclear cells (PBMC), or the  
CC proliferation of endothelial cells), a method for modulating the uptake  
CC of glucose or free fatty acid (FFA) by skeletal muscle cells,  
CC a method for inhibiting the binding of A-peptide to factor VIIA,  
CC or the differentiation of adipocyte cells, a method for detecting the  
CC presence of a tumour in a mammal and an oligonucleotide probe derived  
CC from any of the nucleotide sequences cited above. The nucleic acids and  
CC polypeptides are useful for treating inflammatory diseases, organ  
CC failure, atherosclerosis, cardiac injury, infertility, birth defects,  
CC premature aging, AIDS (acquired immunodeficiency syndrome), cancer, or  
CC diabetic complications. The nucleic acids are useful as hybridisation  
CC probes, in chromosome and gene mapping, and in generating antisense RNA  
CC or DNA. The polypeptides are useful as pharmaceuticals, diagnostics,  
CC biosensors or bioreactors. Both are useful in tissue typing.  
CC The present sequence encodes a PRO protein of the invention.

Search completed: February 15, 2004, 01:20:30  
Job time : 692 secs

QY 515 CATGGCAGACTCTGTAAGAGTGTGTCGACAAATGGGAAGAGCTCCTTGGC---CAGGA 571  
DB 698 GATGGCTCATTTCTGTGAAAATGATGTGATTAAGTGGAGAAAGATTGGCAGCACTCAGGA 757  
QY 572 TTCCCTCTGAGGTCCTTTCAGCAGCTCTCTTGATGAACCTTGACACCATCATGAAGTG 631  
DB 758 CACAAGCGTGAAGTCTATGAGCACATCAACTCGATGTCTCGATATATAATCATGAATG 817  
QY 632 TGCCTTCAGCCATCAGGCGACGATCCAGGTGACAGAAATCTCAGTCTACATACAGGC 691  
DB 818 CGCTTTCAGCAAGAGAACCACTGCCAGACAAACAGACCATGATCCTATGCAAAAGC 877  
QY 692 CATTAGTACCTGAACAACTGGTTTTTCCCGTGTGAGGAATGCTTTCACAGAAATGA 751  
DB 878 CATATTGAACTCAGCAAAATCATATTTCACCGCTGTGACAGTTGTGTATCACAGTGA 937  
QY 752 CACCATCTACAGCCTGAACCTCTGCTGGCCGCTGACACACCGCGCTGCGCAGCTGGCCCA 811  
DB 938 CATTAATTTCAAACTCAGCCCTCAGGGCTACCGCTTCAGAAAGTTAAGCCGAGTGTGAA 997  
QY 812 TCAGCACACAGACCAAGTATCCAACTGAGAGAGCTCAACTACAGAAAGAGGGAGCT 871  
DB 998 TCAGTACACAGATACATAATCCAGAAAGAAAGAAATCCCTCCAGGCTGGGGTAAAGCA 1057  
QY 872 GGAGAAGATCAAGAGAGAGAGGATTTGATTTCTGATATCCTCTCTGGCCAAAT 931  
DB 1058 GGATAACACTCCGAGAGAGAGAGTACCAGGATTTCTGATATTGTCTTCTGCCAAGGA 1117  
QY 932 GGAGATGGGAGCATCTTGTCAAGACAAAGACCTCCGTGCTGAGGTGACACGTTATGTT 991  
DB 1118 TGAAGTGTAGCAGCTTCTCAGATATTGATGTACACTGAAAGTGAAGCAATTCCTGTT 1177  
QY 992 TGAGGGCCACGACACACAGCCAGTGGAGTCTCTGATCCTTAATGCTCTGGCCACACA 1051  
DB 1178 GGCAGGACATGACACTTGGCAGCAGAGCATCTCTGATCCTTAATGCTCTGGCTCTGAA 1237  
QY 1052 CCCCAGCATCAGAGAGGTGCGGAGAGATCCACAGCCTCCTGGGTGATGAGCCTC 1111  
DB 1238 CCCTGAGCATCAAGAGAGATGCGGAGAGAGTCAAGGGCATCCTGGGGATGGGTCTTC 1297  
QY 1112 CATCACTGGAACCACTGAGCCAGATGCCCTACACCATGTGCAATTAAGAGGCACT 1171  
DB 1298 TATCACTTGGGACCACTGGGTGAGATGTCTGTACACCAATGTGCAATCAAGGAGACGTG 1357  
QY 1172 GAGGCTTAACCAAGGTGCGAGGATTTGGCAGAGAGCTCAGCACTCCGTCACCTTCCC 1231  
DB 1358 CCGATTGATTCCTGCAAGTCCCGTCAATTCAGAGATCTCAGCAAGCCACTTACCTTCCC 1417  
QY 1232 TGATGGCGCTCCTTGCCCAAGGATCATGCTCTCTCTCATTTATGACCCTTACCA 1291  
DB 1418 AGATGATGACATGTGCTGACAGGATCACCGGTGTTCTTAGTATTTGGGGTCTTACCA 1477  
QY 1292 CAACCCAAAAGTGTGCCCCAACCCAGAGGTGTTGACCCCTTCCGTTTGCACCGGTTTC 1351  
DB 1478 CAACCTGTCTGTGAAAACCAAGGTCTTGAACCCCTTGAAGTCTCTCAGAGAGAA 1537  
QY 1352 TGCTCAACACAGCCAC-----GCTTCTGCTGCTTCTCAGAGAGATCAAGAACTGCAT 1405  
DB 1538 TTCTGATCAGAGACACCCCTATGCTTACCAATTTCTCAGCTGATCAAGGAACTGCAT 1597  
QY 1406 TGGGAAACAATTGGCATGAACGAGTGAAGGTGGCCACGGCCCTGACCTGCTCGCTT 1465  
DB 1598 TGGGCAAGAGTTTGCATGATGATTAAGTTAACCATGTGCTTGTCTGCTCAGCTT 1657  
QY 1466 TGAGCTGTGCTGATGCCAACAGAGATCCCAATGCCATGACAGACTGTGTGAATC 1525  
DB 1658 CAGAGTACTCAGAGCCCAACAGGCTCTTACTTTCACCAACATTTATCTCAAGCC 1717  
QY 1526 CAAAATGGAATCACTGCTGTCTCAGAGGCTCCCTAACCTT 1569  
DB 1718 CAAGAATGGATGTATTGTGACCTGAAGAACTCTGATGTT 1761



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 15, 2004, 01:20:39 ; Search time 5297 Seconds  
(without alignments)  
11819.583 Million cell updates/sec

Title: US-09-942-429A-6  
Perfect score: 2576

Sequence: 1 gaattccgcagagatccagc.....aaaaaaaaaacggaattc 2576

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlm:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrt:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1012.8	39.3	2084	11	AK002528 Mus muscu
2	974.4	37.8	1201	13	BX340590 BX340590
3	769.4	29.9	854	12	BI759294 BI759294
4	660.8	25.7	1008	13	BX452362 BX452362

5	631.2	24.5	1201	13	BX422210	BX422210
6	626.4	24.3	635	12	BI759809	BI759809
7	626	24.3	701	13	BX097936	BX097936
8	592.6	23.0	781	10	BG533264	BG533264
9	540.2	21.0	3860	11	AK038526	AK038526
10	522.6	20.3	792	12	BG973046	BG973046
11	513.4	19.9	844	12	BI148819	BI148819
12	510.8	19.8	872	10	BF687317	BF687317
13	499.2	19.4	834	12	BI147726	BI147726
14	490.4	19.0	670	14	CB455420	CB455420
15	482.2	18.7	806	12	BI143913	BI143913
16	481.2	18.7	779	12	BI102811	BI102811
17	478.2	18.6	495	14	CB163562	CB163562
18	472	18.3	835	12	BG970447	BG970447
19	470.2	18.3	864	12	BI103261	BI103261
20	466	18.1	737	14	CB952673	CB952673
21	463.4	18.0	864	12	BI219067	BI219067
22	463	18.0	838	9	AI097751	AI097751
23	462.8	18.0	828	9	AI956431	AI956431
24	460.8	17.9	833	12	BG969653	BG969653
25	456.4	17.7	779	14	CB600740	CB600740
26	454.4	17.6	761	10	BF383974	BF383974
27	449.6	17.5	711	14	CB950195	CB950195
28	447	17.4	746	10	BF789061	BF789061
29	444.8	17.3	876	12	BI101160	BI101160
30	439.2	17.0	599	14	CB457963	CB457963
31	435.8	16.9	782	10	BF782104	BF782104
32	433.6	16.8	733	14	CB594270	CB594270
33	433.6	16.8	831	12	BG970019	BG970019
34	430.8	16.7	745	10	BF788511	BF788511
35	430.6	16.7	761	14	CB950710	CB950710
36	428.8	16.6	753	10	BF783236	BF783236
37	428.4	16.6	696	12	BI331167	BI331167
38	426.6	16.6	769	12	BI221979	BI221979
39	424.2	16.5	731	14	CB955255	CB955255
40	424.2	16.5	898	10	BF533291	BF533291
41	424	16.5	649	14	CB417720	CB417720
42	422	16.4	572	10	BF043014	BF043014
43	418.6	16.2	926	10	BF532415	BF532415
44	417.2	16.2	737	14	CB600339	CB600339
45	416.2	16.2	658	12	BI148789	BI148789

ALIGNMENTS

RESULT 1  
AK002528  
LOCUS  
DEFINITION Mus musculus adult male kidney cDNA, RIKEN full-length enriched library, clone:0610011D16 product:cytochrome P450, 4a10, full insert sequence.  
2084 bp mRNA linear HTC 05-DEC-2002

ACCESSION AK002528  
VERSION AK002528.1  
KEYWORDS GI:12832575  
SOURCE HTC; CAP-trapper.  
ORGANISM Mus musculus (house mouse)

REFERENCE  
1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636

REFERENCE  
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
MEDLINE 20499374  
PUBMED 11042159



Qy 429 CTGTTGAATGGGACACATGTTCCAGATCGACGATGCTGATCCCGACCTTCCACTAT 488  
Db 430 TTGTTGAATGACAGCCATGTTCCAGACACAGAGATGCTAACCCGCTTCCACTAT 489  
Qy 489 GACATCTGAAGCCCTATGTTGGGCTCATGGCAGACTCTGTACGATGATGCTGACAA 548  
Db 490 GACATCTGAAGCCCTATGTTGAAGAACATGGCCGACTCCATTCGATGATGCTGACAA 549  
Qy 549 TGGGAAGACTCTTGGCCAGATTTCCCTCTGAGGCTTTCAAGCAGTCTCTGATG 608  
Db 550 TGGGAAGGCTGGCTGATCAAGACTCTCTATAGAGATCTTCAACACATCTCTTAATG 609  
Qy 609 ACCCTGGAACCATCATGAAGTGTGCTTCAAGCCATCAGGCGAGCATCCAGGTGACAG 668  
Db 610 ACCCTGGAACCTGTATGAATGTGCTTCAAGCCACAGGCGAGCTTCAAGTGAATGA 669  
Qy 669 AATTCAGTCTCTACATACAGGCGCAATGACCTGAACACTGTTTTCCTGCTG 728  
Db 670 AATTCAGGACTTACCTCCAGGCGCAATGAGGACTTGAATACCTATTCACTCCGCTG 729  
Qy 729 AGGAATGCTTTACACGAAATGACCATCTACAGCTGACTGCTGCGCGCTGACA 788  
Db 730 AGGAATATCTTTATCATGAATGATACCTACAACTTTCTTCAATGCGCGCTGACC 789  
Qy 789 CACCGGCTGCGACAGTGGCCCATCAGCACACAGACCAAGATGCAACTGAGAGAGCT 848  
Db 790 AAACAAGCTGTCACTTGCCTGATGATCACACAGATGATGATCAAGCTGAGAAAGAT 849  
Qy 849 CAATACAGAAAGAGGAGAGTGAAGATCAAGAGAGAGGCAATTTGATTTCTG 908  
Db 850 CAGTGCAGAGATGAGGAGAGAGTGAAGAAAGATCAAGAAAGAGAGCTTTGATTTCTG 909  
Qy 909 GATATCTCTCTTGGCCAAATGAGATGGAGCATTTGTCAAGACAGACTCTCGT 968  
Db 910 GATATCTTTTATTTGCGAATGAGATGGGACAGCATGTCTGACAAAGACTTAAGT 969  
Qy 969 GCTGAGGTGACACGTTATGTTGAGGGCCACAGACACCAAGCAAGTGGATCTCTG 1028  
Db 970 GCTGAGGTGACACATTCATGTTGAGGGCCATGACACCAAGCAAGTGGATCTCTG 1029  
Qy 1029 ATCTCTATGCTTGGCCACACACCCCAAGCATCAGAGAGGTGCGGAGAGATTCAC 1088  
Db 1030 ATCTCTATGCTTGGCCACACACCTGATCACCACAGAGATGAGAGAGATTCAG 1089  
Qy 1089 AGCTCTCTGGGTGATGAGGCTTCATCACTGGAACCACTGGAACAGATGCCCTACAC 1148  
Db 1090 AGCTCTCTGGGGATGATCTCTATCACTGGGATCACTAGACAGATTCCTCTATACC 1149  
Qy 1149 ACCATGTGATTAAAGAGGCACTGAGGCTTACCCACCGGTGCCAGGATTTGGCAGAG 1208  
Db 1150 ACAATGTGATTAAAGAGGCTTAAGGCTTACCCACCTGTCCAGGCAATTTGAGAGAA 1209  
Qy 1209 CTCAAGCACTCCGTCACCTTCCCTGATGGGCGCTCTTGCCCAAGGTATCATGCTCTC 1268  
Db 1210 CTCAAGTATCTGTCACTTCCCTGATGAGCGCTTTTACCAAAAGGTGTCCAGGTACA 1269  
Qy 1269 CTCTCATTTATGCTTCAACCAACCAAAAGTGTGCGCCCAACCAAGAGTGTTCAC 1328  
Db 1270 CTCTCATTTATGCTTCAACCAACCAAGGTGTGCGCCCAACCAAGAGTGTTCAC 1329  
Qy 1329 CTTTCCGTTTTCACCGGGTCTGCTCAACACAGCAGCTTCTGCTCTCAAGGA 1388  
Db 1330 CTTTCCAGTTTTCACCGAGCTCTCCCGACACAGCCTCATCTGCTCTCAAGGA 1389  
Qy 1389 GGAATCAAGAACTGCTGATGGGAAACAATTGCTGATGAGAGCTGAAGGTGCGCAGG 1448  
Db 1390 GGAATCAAGAACTGCTGATGGGAAACAATTGCTGATGAGAGCTGAAGGTGCTG 1449  
Qy 1449 CTGACCTGCTCGCTTTCAGTGTGCTGCTGATCCCAAGAGATCCCATCCCATTCGA 1508  
Db 1450 CTGACCTGCTCGCTTTCAGTGTGCTGCTGATCCCAAGAGATCCCATCCCATTCGA 1509  
Qy 1509 CGACTTGTGTGAATCCAAAAATGGAATCCACCTGCTCTCAAGAGGCTCC 1560

Db 1510 CGACTTGTCTGAAGTCCAAAAATGGATCTTACCTACATCTCAAGAGCTCC 1561  
RESULT 2  
BX340590  
LOCUS  
DEFINITION  
BX340590 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED  
ACCESSION  
BX340590  
VERSION  
BX340590.1 GI:30335923  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
1 (bases 1 to 1201)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
JOURNAL  
COMMENT  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 947.r for  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DJ003BE03QPI&cluster=947.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paraday Avenue Genoscope sequence ID : CS0DJ003BE03QPI.  
FEATURES  
source  
1..1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DJ003Y106"  
/cell\_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
/cell\_line="JURKAT"  
/clone\_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dr)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT.6 vector. Library was normalized."  
BASE COUNT  
261 a 311 c 326 g 257 t 46 others  
ORIGIN  
Query Match 37.8%; Score 974.4; DB 13; Length 1201;  
Best Local Similarity 94.0%; Pred. No. 1.3e-94;  
Matches 998; Conservative 21; Mismatches 42; Indels 1; Gaps 1;  
Qy 7 CGCAGAGATCCAGCAGGTGCTGACCAATGAGTCTCTGTGCTGAGCCCCAGCAGACTCC 66  
Db 58 CTCAGAGATCCAGCAGGKGTGACCAATGAGTCTCTGTGCTGAGCCCCAGCAGACTCC 117  
Qy 67 TGGGTGATGCTCTGGAATCTCCAGCGGCTCCCTGCTCATTTCTGCTTCTGCTGTA 126  
Db 118 TGGGTGATGCTCTGGAATCTCCAGCGGCTCCCTGCTCATTTCTGCTTCTGCTGTA 177  
Qy 127 TCAAGGAGTTCAGCTCTACCTGACAGGAGTGGCTGCTCAAGCCCTCCAGCAGTTCC 186  
Db 178 TCAAGGAGTTCAGCTCTACCTGACAGGAGTGGCTGCTCAAGCCCTCCAGCAGTTCC 237  
Qy 187 CGTGCCCTCCCTCCCACTGGCTCTTGGGCAATCCAGAGAGCTTCAACAGAGCAGAGC 246  
Db 238 CGTGCCCTCCCTCCCACTGGCTCTTGGGCAATCCAGAGAGCTTCAACAGAGCAGAGC 297  
Qy 247 TACAAGGATTCAGAAATGGGTGAGACATTCACCAAGTGCCTGTCTCATTTGGCTATGG 306  
Db 298 TACAAGGATTCAGAAATGGGTGAGACATTCACCAAGTGCCTGTCTCATTTGGCTATGG 357

Oy		307	GAGGCMAAGTTCGTGCTCCAGCTCTATGACCTTGACTATAATGAAGTGATTCTGGGAGAT	366
Db		358	GAGGCMAAGTTCGTGCTCCAGCTCTATGACCTTGACTATAATGARGGTGATTCTGGGAGAT	417
Oy		367	CAGACC CGAAATCCC ATGGTTCTCCTACAGATTCTGGCTCATGATTGGTAGCGGCTTGC	426
Db		418	CAGACC CGAAATCCC ATGGTTCTCCTACAGATTCTGGCTCATGATTGGTAGCGGCTTGC	477
Oy		427	TCCGTGTTGAATGGG CAGACATGGTTCAGCATCGA CCGATGCTGACCCCA GCTTCCA CT	486
Db		478	TCCTGTTGAATGGG CAGACATGGTTCAGCATCGA CCGATGCTGACCCCA GCTTCCA CT	537
Oy		487	ATGACATCTTGAA GCCCTATATGTGGGCTCATGGCAGA CTCTGTACGAGTGA TGTCTGACA	546
Db		538	ATGACATCTTGAA GCCCTATATGTGGGCTCATGGCAGA CTCTGTACGAGTGA TGTCTGACA	597
Oy		547	AATGGGAAGAGCT CCTTGGCCAGGATTCCTCCCTTGAGGCTTT CAGCACGTC TCCTTGA	606
Db		598	AATGGGAAGAGCT CCTTGGCCAGGATTCCTCCCTTGAGGCTTT CAGCACGTC TCCTTGA	657
Oy		607	TGACCTTGACACAC CATATGAAGTGTGCC TTCA GCCATCAGGGCAGCATCCAGGTGACA	666
Db		658	TGACCTTGACACAC CATATGAAGTGTGCC TTCA GCCATCAGGGCAGCATCCAGGTGACA	717
Oy		667	GGAATTCCTCAGT CCTTACATACAGGCCATTAGTACCTTGAACA CAACCTGTTTTTCCCG TG	726
Db		718	GGAATTCCTCAGT CCTTACATACAGGCCATTAGTACCTTGAACA CAACCTGTTTTTCCCG TG	777
Oy		727	TGAGGAATGCC TTTCA CCAAGATGACACCATCTACAGCTGACCTCTGCTGGCCGCTGGA	786
Db		778	TGAGGAATGCC TTTCA CCAAGATGACACCATCTACAGCTGACCTCTGCTGGCCGCTGGA	837
Oy		787	CACACCGGCTGCG CAGCTGGCCCATCAGCACACAGCA CCAAGTGATCCAACTGAGGAAG	846
Db		838	CACACCGGCTGCG CAGCTGGCCCATCAGCACACAGCA CCAAGTGATCCAACTGAGGAAG	897
Oy		847	CTCAACTACAGA AGAGGGGAGCTGAGAGAATCAAGA GAAGAAGGCATTGTGATTTC	906
Db		898	CTCAACTACAGA AGAGGGGAGCTGAGAGAATCAAGA GAAGAAGGCATTGTGATTTC	957
Oy		907	TGATATCTCTCT TTTGGCCAAAATGAGAA TGGAGCATTTGT CAGACAAGGACTCC	966
Db		958	TGATATCTCTCT TTTGGCCAAAATGAGAA TGGAGCATTTGT CAGACAAGGACTCC	1017
Oy		967	GTGCTGAGGTGACA CGTTCATATGTTGAGGGCCAGACA CACAGCCAGTCAGGATCTCT	1026
Db		1018	GTGCTGAGGTGACA CGTTCATATGTTGAGGGCCAGACA CACAGCCAGTCAGGATCTCT	1076
Oy		1027	GGATCTCTATGCT CTGGCCACACACCCCAAGCATCAGGAGA	1068
Db		1077	TCTTATCTTKGCMWMMCMAGCATAGGGGTGCGGRGR	1118
RESULT 3				
LOCUS	BI759294	854 bp	mRNA	linear EST 25-SBP-2001
DEFINITION	603042936P1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183289 5',			
ACCESSION	BI759294			
VERSION	BI759294.1	GI:15750872		
KEYWORDS	EST.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
TITLE	NIH-MGC http://mgc.nci.nih.gov/.			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT	Unpublished			
	Contact: Robert Strausberg, Ph.D.			
	Email: cgabbs-r@mail.nih.gov			
	Tissue Procurement: Life Technologies, Inc.			
	cDNA Library Preparation: Life Technologies, Inc.			

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:			
http://image.llnl.gov			
Plate: LLAM1457 row: h column: 10			
High quality sequence stop: 830.			
Location/Qualifiers			
1. 854			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="IMAGE:5183289"			
/lab_host="DH10B"			
/clone_lib="NIH_MGC_116"			
/note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH MGC Library."			
BASE COUNT	187	a	252 c 223 g 192 t
ORIGIN			
Query Match	29.9%	Score 769.4;	DB 12; Length 854;
Best Local Similarity	98.5%;	Pred. No. 8.2e-73;	
Matches 840; Conservative	0; Mismatches 6;	Indels 7;	Gaps 6;
QY	7	CGCAGAGATCCAGCAGGTGCTGCACCATGAGTGTCTGTGCTGAGCCCCAGCAGACTCC	66
Db	6	CTCAGAGATCCAGCAGGTGCTGCACCATGAGTGTCTGTGCTGAGCCCCAGCAGACTCC	65
QY	67	TGGGTGATGTTCTGGAATCTCCAGCGGCTCCCTGCTCATTTCTGTTCTGCTGTA	126
Db	66	TGGGTGATGTTCTGGAATCTCCAGCGGCTCCCTGCTCATTTCTGTTCTGCTGTA	125
QY	127	TCAAGCAGTTCAAGTCTTAAGTCTGACAGGAGGAGTGGTCTCAAAAGCCCTCCAGCAGTTCC	186
Db	126	TCAAGCAGTTCAAGTCTTAAGTCTGACAGGAGGAGTGGTCTCAAAAGCCCTCCAGCAGTTCC	185
QY	187	CGTGCCCTCCCTCCCACTGGCTTTGGGACATCCAGAGCTCCAACAGGACGAGGC	246
Db	186	CGTGCCCTCCCTCCCACTGGCTTTGGGACATCCAGAGCTCCAACAGGACGAGGC	245
QY	247	TACAACGATTCAAGAAATGGGTGAGACATTTCCCAAGTGTCTCTATTGGCTATGGG	306
Db	246	TACAACGATTCAAGAAATGGGTGAGACATTTCCCAAGTGTCTCTATTGGCTATGGG	305
QY	307	GAGGCAAGTTGCTGTCCAGCTTATGACCCCTGACTATATGAAGGTGATTCTGGGAGAT	366
Db	306	GAGGCAAGTTGCTGTCCAGCTTATGACCCCTGACTATATGAAGGTGATTCTGGGAGAT	365
QY	367	CAGACCCGAAATCCCAATGGTCTTACAGATTCTGGCTCCATGATTTGGGTACGGCTTGC	426
Db	366	CAGACCCGAAATCCCAATGGTCTTACAGATTCTGGCTCCATGATTTGGGTACGGCTTGC	425
QY	427	TCTGTGTAATGGGACAGATGGTTCAGCATGACGAGTGTGACCCCAAGCTTCCACT	486
Db	426	TCTGTGTAATGGGACAGATGGTTCAGCATGACGAGTGTGACCCCAAGCTTCCACT	485
QY	487	ATGACATCTGAAGCCCTATGTGGGGCTCATGGCAGACTCTGTACGAGTATGCTGACA	546
Db	486	ATGACATCTGAAGCCCTATGTGGGGCTCATGGCAGACTCTGTACGAGTATGCTGACA	545
QY	547	AATGGGAAGAGCTCCTTTGGCCAGGATTCCCTCTGAGGCTTTACGACGCTCTCTTGA	606
Db	546	AATGGGAAGAGCTCCTTTGGCCAGGATTCCCTCTGAGGCTTTACGACGCTCTCTTGA	605





JOURNAL COMMENT

Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 947.r For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DM002AB01QPI&cluster=947.r. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS0DM002AB01QPI.  
Location/Qualifiers  
1.1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DM002YC01"  
/tissue\_type="FETAL LIVER"  
/dev\_stage="fetal"  
/clone\_lib="Homo sapiens FETAL LIVER"  
/note="Organ: liver; Vector: pCMVSPORT 6; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-strand cDNA was digested with Not I and  
cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
vector. Library was not normalized."

FEATURES source

BASE COUNT 275 a 306 c 307 g 269 t 44 others  
ORIGIN

Query Match 24.5%; Score 631.2; DB 13; Length 1201;  
Best Local Similarity 98.8%; Pred. No. 2.8e-58;  
Matches 656; Conservative 1; Mismatches 4; Indels 3; Gaps 2;

OY 7 CGCAGAGATCCAGAGTGTGTCACCATGAGTGTCTGTGTGAGCCCGACAGACT-- 64  
DB 70 CTCAGAGATCCAGAGTGTGTCACCATGAGTGTCTGTGTGAGCCCGACAGACTCA 129  
OY 65 CCTGGGTGATGTCCTG-GAATCCTCCAGCGGCTCCCTGCTCATTTGCTTGTGCTGC 123  
DB 130 CTGAGGTGATGTCCTGAGAAATCTCCAGCGGCTCCCTGCTCATTTGCTTGTGCTGC 189  
OY 124 TGATCAAGGCAAGTCAAGCTCTACCTGACAGGCAAGTGGCTCAAGCCCTCCAGCAGT 183  
DB 190 TGATCAAGGCAAGTCAAGCTCTACCTGACAGGCAAGTGGCTCAAGCCCTCCAGCAGT 249  
OY 184 TCCCGTCCCTCCCTCCCACTGGCTCTTGGGCAATCCAGAGCTCCACAGAGCCAGG 243  
DB 250 TCCCGTCCCTCCCTCCCACTGGCTCTTGGGCAATCCAGAGCTCCACAGAGCCAGG 309  
OY 244 AGCTACAACGATTCAGAAATGGGTGAGAGACATTCCTCAAGTGCCTGCTCATTTGGCTAT 303  
DB 310 AGCTACAACGATTCAGAAATGGGTGAGAGACATTCCTCAAGTGCCTGCTCATTTGGCTAT 369  
OY 304 GGGGAGGCAAGTTCGTGTCAGCTCTATGACCTGACTATATGAAGGTATTTGGGGA 363  
DB 370 GGGGAGGCAAGTTCGTGTCAGCTCTATGACCTGACTATATGAAGGTATTTGGGGA 429  
OY 364 GATCAGACCCGAATCCATGTTCTCTACAGATTCCTGGCTCCATGATGGGTACGGCT 423  
DB 430 GATCAGACCCGAATCCATGTTCTCTACAGATTCCTGGCTCCATGATGGGTACGGCT 489  
OY 424 TGCTCTGTGTAATGGGCAAGATGTTCCAGCATGACGGATGCTGACCCAGCCTTCC 483  
DB 490 TGCTCTGTGTAATGGGCAAGATGTTCCAGCATGACGGATGCTGACCCAGCCTTCC 549  
OY 484 ACTATGACATCTGAAGCCTATATGTGGGGCTCATGGCAGACTCTGTACGAGTATGCTGG 543  
DB 550 ACTATGACATCTGAAGCCTATATGTGGGGCTCATGGCAGACTCTGTACGAGTATGCTGG 609  
OY 544 ACAATGGAGAGAGCTCTTGGCCAGGATTCCTCTGTGAGGTCTTTCAAGCAGTCTCCT 603  
DB 610 ACAATGGAGAGAGCTCTTGGCCAGGATTCCTCTGTGAGGTCTTTCAAGCAGTCTCCT 669

OY 604 TGATGACCTTGACACCATCATGAGTGTGCTTACGCCATCAGGCGACATCCAGGTGG 663  
DB 670 TGATGACCTTGACACCATCATGAGTGTGCTTACGCCATCAGGCGACATCCAGGTGG 729  
OY 664 ACAG 667  
DB 730 ACAG 733

RESULT 6  
BI759809  
LOCUS  
DEFINITION  
BI759809 635 bp mRNA linear EST 25-SEP-2001  
603045693F1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5186015 5',  
mRNA sequence.  
BI759809  
BI759809 1 GI:15751387  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LAM1464 row: 1 column: 24  
High quality sequence stop: 632.  
Location/Qualifiers

FEATURES source

1.635  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5186015"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_116"  
/note="Organ: pooled colon, kidney, stomach; Vector:  
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA  
source anonymous pool of 3 colons, age 26 yo male, 49 yo  
female, 71 yo male colon; 46 yo male kidney, and pool of 2  
stomachs, 62 yo male and 70 yo female. Library is  
oligo-dT primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.4 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
023. Note: this is a NIH MGC Library."

BASE COUNT 130 a 186 c 170 g 148 t 1 others  
ORIGIN

Query Match 24.3%; Score 626.4; DB 12; Length 635;  
Best Local Similarity 99.8%; Pred. No. 1.3e-57;  
Matches 627; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 7 CGCAGAGATCCAGAGTGTGTCACCATGAGTGTCTGTGTGAGCCCGACAGACTCC 66  
DB 7 CTGAGAGATCCAGAGTGTGTCACCATGAGTGTCTGTGTGAGCCCGACAGACTCC 66  
OY 67 TGGGTGATGTCCTGGAATCTCCAGCGGCTCCCTGCTCATTTGCTTGTGCTGTA 126  
DB 67 TGGGTGATGTCCTGGAATCTCCAGCGGCTCCCTGCTCATTTGCTTGTGCTGTA 126  
OY 127 TCAAGCAGTTCAAGCTTACCTGACAGGCAAGTGGCTGCTCAAGCCCTCCAGCAGTTC 186  
DB 127 TCAAGCAGTTCAAGCTTACCTGACAGGCAAGTGGCTGCTCAAGCCCTCCAGCAGTTC 186

Db	127	TCAAGCAGATTCAAGCTCTACCTGCAAGGCAAGTGGCTGCTCAAAAGCCCTCCAGCAGATTCC	186
OY	187	CGTGGCCTTCCCTCCCACTGGCTCTTGGGCACATCCAGAGCTCCAAAGGACCAAGAGC	246
Db	187	CGTGCCCTCCCTCCCACTGGCTCTTGGGCACATCCAGAGCTCCAAAGGACCAAGAGC	246
OY	247	TACAAACGATTCAGAAATGGGTGAGACATTCCCAAGTGCCTGCTCAATGGCTATGGG	306
Db	247	TACAAACGATTCAGAAATGGGTGAGACATTCCCAAGTGCCTGCTCAATGGCTATGGG	306
OY	307	GAGGCAAAAGTTCGTGTCCAGCTCTATGACCCCTGACTATATGAAGTGAATCTGGGAGAT	366
Db	307	GAGGCAAAAGTTCGTGTCCAGCTCTATGACCCCTGACTATATGAAGTGAATCTGGGAGAT	366
OY	367	CAGACCCGAAATCCCATGTTCTCTACAGATTCTCGGCTCCATGGAATGGGTACGGCTTGC	426
Db	367	CAGACCCGAAATCCCATGTTCTCTACAGATTCTCGGCTCCATGGAATGGGTACGGCTTGC	426
OY	427	TCCTGTGAATGGGCAGACATGTTCCAGCATGACGAGTCTGAACCCAGCCTTCCACT	486
Db	427	TCCTGTGAATGGGCAGACATGTTCCAGCATGACGAGTCTGAACCCAGCCTTCCACT	486
OY	487	ATGACATCTGAAGCCCTATGTGGGGCTCATGGCAGACTCTGTACAGATGCTTGAGCA	546
Db	487	ATGACATCTGAAGCCCTATGTGGGGCTCATGGCAGACTCTGTACAGATGCTTGAGCA	546
OY	547	AATGGGAAGAGCTCCTTGGGCCAGGATTCCCTCTGGAAGTCTTTCAAGCAGTCTCCTTGA	606
Db	547	AATGGGAAGAGCTCCTTGGGCCAGGATTCCCTCTGGAAGTCTTTCAAGCAGTCTCCTTGA	606
OY	607	TGACCCCTGGAACCATCATGAAGTGTGC	634
Db	607	TGACCCCTGGAACCATCATGAAGTGTGC	634

RESULT 7					
BX097936					
LOCUS					
DEFINITION					
BX097936	701 bp	mrna	linear	EST 04-FEB-2003	
BX097936	Soares fetal liver spleen	INFLS	Homo sapiens	CDNA clone	
IMAGE998A11111	; IMAGE:120466,	mrna	sequence.		

ACCESSION	BX097936
VERSION	BX097936.1
KEYWORDS	EST.

SOURCE ORGANISM	Homo sapiens (human)
	Homo sapiens

REFERENCE  
1 (bases 1 to 701)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
1 (bases 1 to 701)  
Ebert, L., Hell, O., Hennig, S., Neubert, P., Patsch, B., Peters, M.,  
Radelof, U., Schneider, D. and Korn, B.

TITLE	Human Unigeneset - R2PD3
JOURNAL	Unpublished

**COMMENT**      **Contact: Ina Rolfs**

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

RZPD; IMAGP998A11111

RZPD LIB; I.M.A.G.E. cDNA Clone Collection;  
Human UniGeneSet - RZPD3 (RZPD LIB No.972)

<http://www.rzpd.de/CloneCards/cgi->

bin/showlib.pl.cgi?response?libNo=972 Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Heubnerweg 6, D-14059 Berlin, Germany

Tel: +49 30 32639 101  
Fax: +49 30 32639 111

**www.rzpd.de**

This clone is available royalty-free from RZPD; contact RZPD ([clone@rzpd.de](mailto:clone@rzpd.de)) for further information. Seq primer: M13r, Primer sequence: TTTCACACAGGAACAGCTATGAC.

### FEATURES

#### Location/Qualifiers

**Source**

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"
```

```

BASE COUNT      136 a
ORIGIN
/clone="IMAGp998A11111 ; IMAGE:120466"
/sex="male"
/dev stage="20 week post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen 1NFLS"
/note="Organ: Liver and Spleen; Vector: pTT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGAGAGATTAATTAAGATCTTTTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia)', digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pTT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
220 c 151 g 186 t 8 others

```

Query Match	24.3%;	Score 626;	DB 13;	Length 701;
Best Local Similarity	97.9%;	Pred. No. 1.4e-57;		
Matches 652; Conservative	0;	Mismatches 11;	Indels 3;	Gaps 2;

[illegible][illegible]

```

Oy      1579 AGGACCAAGCTTTGAGGGGCTCCACCTGCCGCTCTGTCTTCTGTACCCCCGGCTTCTGTCCC 1638
          |||||
Db      156  AGGACCAAGCTTTGAGGGTCTCCACCTGCCGCTCTGTCTTCTGTACCCCCGGCTTCTGTCCC 215
          |||||

```

QY 1639 CTTCCTGTCTGCCATATCCTGTTTCTGTCTGCCACCTCCCTTCTTCCCACTGGCT 1698  
|||  
Db 216 CTTCCGTCTGCCCATATCCTGTTTTCTGTCTGCCACCTTCCCTTCTTCCCACTGGCT 275

QY 1699 GCTGTCGCCAGTGTGCCTGACCCCTTCTCTCTCAACTTTCTCCAGGCTCCCTACTGCT 1758  
|||  
Db 276 GCTGTCCGCCAGTGTGCCTGACCCCTTCTCTCTCAACTTTCTCCAGGCTCCCTACTGCT 335

QY 1759 TGTCTACCTGTCCTTACCACCACTGTATCTCTTTGGGAAGAAAGCTGAGTGTTGGAG 1818  
| | | | | | | | | | | | | | | | | | | | | |  
Db 336 TGTCTACTGTCCTTACCACCACTGTATCTCTTTGGGAAGAAAGCTGAGTGTTGGAG 395

QY 1819 AAGTGAGGCCGAGTTGCATGTCTGACATAATGTAAGAAGCTTGAATCATGTCCAGG 1878  
 |||||  
 Db 396 AAGCTGAGGCCGAGCTTGCATGTCTGACATAATGTAAGAAGCTTGAATCATGTCCAGG 455

Qy 1879 ATCCAGGCTTAAACCCCTGTGGCCTTTGGACAACCAAGCTCTGTGCTGAAGGGTGA 193  
|||  
Db 456 ATCCAGGCTTAAACCCCTGTGGCCTTTGGACAACCAAGCTCTGTGCTGAAGGGTGA 515

1939 AGGCTACCCCTGACGACCATTAATCTAAGCCCCGGGGCATATAAACCCCTCGTGGCTTGATA 1938  
 516 AGGCTACCCCTGACGACCATTAATCTAAGCCCCGGGGCATATAAACCCCTCGTGGCTTGATA 575

by	1999	GAAATCAGGGCTGCGGCCTGGATGTGTGCACTGCGGCTCCTTGCTTC	635
Db	576	GAAATCAGGGCTGCGGCCTGGATGTGTGCACTGCGGCTCCTTGCTTC	

[illegible]

2115	ACACAGA	2127
2116		
2117		
2118		
2119		
2120		
2121		
2122		
2123		
2124		
2125		
2126		
2127		
2128		
2129		
2130		
2131		
2132		
2133		
2134		
2135		
2136		
2137		
2138		
2139		
2140		
2141		
2142		
2143		
2144		
2145		
2146		
2147		
2148		
2149		
2150		
2151		
2152		
2153		
2154		
2155		
2156		
2157		
2158		
2159		
2160		
2161		
2162		
2163		
2164		
2165		
2166		
2167		
2168		
2169		
2170		
2171		
2172		
2173		
2174		
2175		
2176		
2177		
2178		
2179		
2180		
2181		
2182		
2183		
2184		
2185		
2186		
2187		
2188		
2189		
2190		
2191		
2192		
2193		
2194		
2195		
2196		
2197		
2198		
2199		
2200		
2201		
2202		
2203		
2204		
2205		
2206		
2207		
2208		
2209		
2210		
2211		
2212		
2213		
2214		
2215		
2216		
2217		
2218		
2219		
2220		
2221		
2222		
2223		
2224		
2225		
2226		
2227		
2228		
2229		
2230		
2231		
2232		
2233		
2234		
2235		
2236		
2237		
2238		
2239		
2240		
2241		
2242		
2243		
2244		
2245		
2246		
2247		
2248		
2249		
2250		
2251		
2252		
2253		
2254		
2255		
2256		
2257		
2258		
2259		
2260		
2261		
2262		
2263		
2264		
2265		
2266		
2267		
2268		
2269		
2270		
2271		
2272		
2273		
2274		
2275		
2276		
2277		
2278		
2279		
2280		
2281		
2282		
2283		
2284		
2285		

RESULT 8					
BG533264					
LOCUS	BG533264		781 bp	mRNA	linear
					EST 03-APR-2001

DEFINITION 601860630F2 NIH\_MGC\_76 Homo sapiens cDNA clone IMAGE:4072062 5', mRNA sequence.  
ACCESSION BG533264  
VERSION BG533264.1 GI:13524804  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 781)  
NIH-MGC http://mgs.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
AUTHORS Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: CLONETECH Laboratories, Inc.  
CDNA Library Preparation: CLONETECH Laboratories, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.lnl.gov  
Plate: LNC918 row: k column: 07  
High quality sequence stop: 766.  
Location/Qualifiers  
1. 781  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4072062"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH MGC 76"  
/note="Organ: liver; Vector: pDNR-LIB (Clontech); Site\_1: SfiI (ggcgctcgcc); Site\_2: SfiI (ggccattatggc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.85 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH MGC Library."  
BASE COUNT 168 a 227 c 204 g 182 t  
ORIGIN  
Query Match 23.0%; Score 592.6; DB 10; Length 781;  
Best Local Similarity 96.0%; Pred. No. 4.5e-54;  
Matches 629; Conservative 0; Mismatches 24; Indels 2; Gaps 2;  
QY 7 CGCAGAGATCCAGAGGTGTGTGACACCATGAGTCTGTGTGTGAGCCCGACAGACTCC 66  
DB 19 CTGAGAGATCCAGAGGTGTGTGACACCATGAGTCTGTGTGTGAGCCCGACAGACTCC 78  
QY 67 TGGGTGATCTCTGGAATCTCCAGCGGCTCCCTGCTCATTTGCTTCTGTGCTGA 126  
DB 79 TGGGTGATCTCTGGAATCTCCAGCGGCTCCCTGCTCATTTGCTTCTGTGCTGA 138  
QY 127 TCAAGGCGATTCAGCTCTACCTGCACAGGAGTGTGCTCTCAAGCCCTCCAGAGTTCC 186  
DB 139 TCAAGGCGATTCAGCTCTACCTGCACAGGAGTGTGCTCTCAAGCCCTCCAGAGTTCC 198  
QY 187 CGTGCCCTCCCTCCACTGGCTCTTGGGACATCCAGAGCTCCAAGACAGACAGAGC 246  
DB 199 CGTGCCCTCCCTCCACTGGCTCTTGGGACATCCAGAGCTCCAAGACAGAGAGC 258  
QY 247 TACAACGATTCAGAAATGGGTGAGAGACTCCCAAGTGCCTGCTCATTTGGCTATGGG 306  
DB 259 TACAACGATTCAGAAATGGGTGAGAGACTCCCAAGTGCCTGCTCATTTGGCTATGGG 318  
QY 307 GAGGCAAGTTCGTGTCCAG-CTCTATGACCTGACTATATGAGGTGATTCTGGGAGA 365  
DB 319 GAGGCAAGTTCGTGTCCAGACTCTATGACCTGACTATATGAGGTGATTCTGGGAGA 378

QY 366 TCAGACCCGAAATCCCATGTTCTCTACAGATTCTCGTCCATGATT-GGGTACGGCTT 424  
DB 379 TCAGACCCGAAATCCCATGTTCTCTACAGATTCTCGTCCATGATTGGGTACGGCTT 438  
QY 425 GCTCCTGTTGATGGGACAGACATGTTCCAGCATGACGAGATGCTGACCCGACCTTCCA 484  
DB 439 GCTCCTGTTGATGGGACAGACATGTTCCAGCATGACGAGATGCTGACCCGACCTTCCA 498  
QY 485 CTATGACATCTGAAGCCCTATGTGGGCTCATGCGACACTCTGTACGAGTGTGTTGA 544  
DB 499 CTATGACATCTGAAGCCCTATGTGGGCTCATGCGACACTCTGTACGAGTGTGTTGA 558  
QY 545 CAAATGGGAAGCTCTCTGGCCAGAGATTCCCTCTGAGGTCTTTACAGCAGTCTCCTT 604  
DB 559 CAAATGGGAAGCTCTCTGGCCAGAGATTCCCTCTGAGGTCTTTACAGCAGTCTCCTT 618  
QY 605 GATGACCTTGACACCATCATGAGTGTGCTTCAGCCATCAGGCGACATCCAG 659  
DB 619 GATGACCTTGACACCATCATGAGTGTGCTTCAGGCAATTGGCAGAGAGCTCAG 673  
RESULT 9  
AK038526 3860 bp mRNA linear HTC 05-DEC-2002  
LOCUS Mus musculus adult male hypothalamus cDNA, RIKEN full-length  
DEFINITION enriched library, clone:A230025G20 product:similar to SIMILAR TO  
CYTOCHROME P450, 4A family [Mus musculus], full insert sequence.  
ACCESSION AK038526 1 GI:26332620  
VERSION AK038526 1  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
JOURNAL 99279253  
MEDLINE 10349636  
PUBMED  
REFERENCE 2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
JOURNAL 20499374  
MEDLINE 11042159  
PUBMED  
REFERENCE 3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, B., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
JOURNAL 20530913  
MEDLINE 11076861  
PUBMED  
REFERENCE 4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, K., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, S., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stabli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,



Fletcher, C., Fujita, M., Gariboldi, M., Guetincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kotsuki, S. and Hayashizaki, Y.

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

TITLE

The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)  
6 (bases 1 to 3860)

JOURNAL  
REFERENCE  
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE  
JOURNAL

Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physiological and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
location/Qualifiers

FEATURES

source

CDS

1. 3860  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="PANTOM DB:A230025G20"  
/db\_xref="taxon:10090"  
/clone="A230025G20"  
/sex="male"  
/tissue\_type="hypothalamus"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
1. 1500  
/note="unnamed protein product; putative similar to SIMILAR TO CYTOCHROME P450, 4A family [Mus musculus] (SPTRIQ91W5, evidence: FASTV, 52.3%id, 96.2%length, match=1473)"  
/codon\_start=1  
/protein\_id="BAC30028.1"  
/db\_xref="GI:26332621"  
/translation="GWARPLHALVFCIALVLMQMKLYLRQRLLRDLSPBPBPAH WLGHQKFLQEDNMTLDELVKHPCAFPCWVGPFQAFYIYDPDYAKIFLSRTDKM OYLHQULTPCIGRLNLDGRFQHRCLLTPAFQDILKPCVDVMAHSVYKMLDKMB KMTTQETTIEVEFHINLMTLDIMKCAFGOETNCOINGTYESYVKAIFBGEIISR LYNFMHHDIIFKLSPKGFQBLGVKHQYTKETIQDRKILKNQVKDDTQTSQIF LDIIVLSAQABDRAFSDADLRRAEVNTFMWAGHDASASISWLLYCLALNPRHQDRCT

polyA_signal		EIRSIIGDSSITWEQDEMSTTMCIKETLRILRPVPSISRELSKPLTPDGHSLPA	
polyA_site		GMTVVISIWGLHNPVWMDPKVFDPLRFTKENSQDRHPCALFPSSGPRNCIGQOFA	
BASE COUNT		1004 a 883 c 866 g 1107 t	
ORIGIN		MELKVAIALILHLPQVADLTRPPAFSSHTVLRPRGIIYLHLKLEEC"	
Query Match		21.0%; Score 540.2; DB 11; Length 3860;	
Best Local Similarity		62.1%; Pred. No. 5.9e-49;	
Matches 927; Conservative		0; Mismatches 548; Indels 18; Gaps 4;	
QY	86 CCTCAAGCGCCTCCCTGCTCATCTCTGCTTCTGCTGATCAAGCAGTTCAGCTCTA	145	
DB	15 CTGCACTGCTTGTGTTCTGCTGGCGCTGCTGATGAGCCATGAGCTCTA	74	
QY	146 CTTGACATGAGCATGCTGCTCAAAGCCCTCCAGCAGTTCGCCGTCCTCCCACTG	205	
DB	75 CTTGCCGAGACAGCAGCTGCTGCCGCACTGAGCCCTTCCCGGCGCTCCGCCACTG	134	
QY	206 GCTCTTGGGCAATCCAGAGCTTCAACAGACAGCAGGAGCTTCAACAGGATTCAGAATG	265	
DB	135 GCTCTGGACA--CCAGAGTTCTTCAGAGAGTATATATGAGACGCTGATGAGAT	191	
QY	266 GGTGAGACATTCACAGTGCCTGCTCATTTGGCTATGGGAGGCAAGTTGTTCCA	325	
DB	192 TGTCAAAAGACACCTGTGCTTCCCTGCTGGTAGGGCCCTTCCAGCATTTTCTA	251	
QY	326 GCTCTATGACCTGACTATATGAGGTATTTCTGGGAGATCAACCCGAATCCCATGG	385	
DB	252 CATCTATGACCCAGACTATGCGAAGATATTCTGAGCAGACAGACCCAAAGATGACATA	311	
QY	386 TTCCTAGATTCCTGCTGCTCCATGATGGGTACGGCTTCTCTTGAATGGCAGAC	445	
DB	312 TCTGACGAGCTCTGACTCCATGCAATTTGGAACGAGGACTCTGAAATTTAGATGGCCAG	371	
QY	446 ATGTTGACGATGACGAGTCTGACCCGCTTCCACTATGATCTCTGAAGCCCTA	505	
DB	372 GTGTTGACACACCGCTGCTCTTAATCTGCTGATTCATCAGACATCTTAACCATG	431	
QY	506 TGTGGGCTCATGACAGCTCTGAGATGATGCTGACAAATGGGAAGCTCCTTGG	565	
DB	432 TGTGACAGATGGCCCACTGTGAAGATGATGCTGATTAATGGAGAAGATGTGAC	491	
QY	566 C--CAGATTCCCTTGAGGTCTTTCAGCAGCTCTCTGATGACCTTGACACCAT	622	
DB	492 CACTCAGGAACAACCATGAGGTTTGAACACATCACTGATGACCTGACATAT	551	
QY	623 CATGAAGTGTGCTTCAAGCATCAGGAGCATCCAGGTGACAGAAATTTCTAGTCTA	682	
DB	552 AATGAATGTGCTTTGGCCAGGACCACTGCCAGATAAAGGCACTATGAGTCTTA	611	
QY	683 CATACAGGCAATGATGACCTGAACAACCTGTTTTCCTGCTGAGGAATGCTTTCA	742	
DB	612 CGTGAAGCAACATTTGAAGTGTGAATCATATCTTCTGCTGTAATTTCTGCA	671	
QY	743 CCAGATGACACATCTACAGCTGACCTCTGCTGCGCTGACACACCGCGCTGCCA	802	
DB	672 TCATCATGACATAATTTCAAACTCAAGCTTAAGGCGCACTGCTTCCAGAGTTAGGCAA	731	
QY	803 GCTGGCCATCAGACACAGACCAAGTATCCAACTGAGGAAGGCTCAACTACAGAAGA	862	
DB	732 AGTATACATCAATAACAGAAAGATAATCCAGACAGAAAGAAATCTCAAGATCA	791	
QY	863 GGGGAGCTGAGAAAGATCAAGAGAGAGGCAATTTGATTTCTGATATCTCTCTT	922	
DB	792 AGTAAGCAGG-----TGACACTCAGACGCTCAAAATTTTCTGATATTTCTTTC	845	
QY	923 GGCCTAATGAGATGAGGAGCATCTTGTGACAAAGGACCTCGTGTGAGGTGACAC	982	
DB	846 TGCACAGGCTGAGATGAGACGAGCTTCTCAGATGCTGACCTTGGGCTGAGGTGACAC	905	

Qy	983	GTTCATGTTGAGGCCACGACACCAAGCCAGTGGGATCTCTGATCTCTTAAGTCT	1042
Db	906	CTTCATGTGGCAGAGCATGATGCTCTGACGTAAAGATCTCTGGCTCTTAACTGCT	965
Qy	1043	GGCCACACACCCCAAGCATCAGAGAGGTGCCGGAGAGATCCACAGCTCTGGGTGA	1102
Db	966	GGCTCTAATCCCGAGCATCAGACAGATGCCGGAAGAGATCAGAGCATCTGGGAGA	1025
Qy	1103	TGAGCCTCCATCACTTGAACCACTGACCAAGATGCCCTAACCAACCATGTGATTA	1162
Db	1026	TGGGCTTCCATCACTTGGGAAACAGCTGGATGAGATGTCAACACCAATGTGCATCA	1085
Qy	1163	GGAGGCACTGAGGCTCTAACCAACGGGTGCCAGGCATTGGCAGAGAGCTCAGACTCCGT	1222
Db	1086	GGAGACACTCCGCTTGATTCCTCTGTCCCATTCATCTCCAGGAGCTCAGCAACCCCT	1145
Qy	1223	CACCTTCCCTGATGGGCGCTCTTGCCCAAGGATCATAGTCTCTCCATTAAAG	1282
Db	1146	TACCTCCAGATGGAACACTCACTGCTGCAGGTATGACTGTGTTCTTAGATTTGGGG	1205
Qy	1283	CCTTCAACCAACCCCAAAAGTGTGGCCCAACCCAGAGGTGTTGAACCTTTCGTTTGC	1342
Db	1206	TCTCCACCACAAACCTGTGTCTGGAACGACCCAAAGGTCTTGAACCCCTTGAGATTAC	1265
Qy	1343	ACCGGTTCTGCTCAACACAGCCAC-----GCTTCTCTGCCCCCTCTCAGAGGATCAAG	1396
Db	1266	TAAAGAAATTCTGATCAGAGACACCCCTGTGCTTCTTAACATTCTCAGTGTGTCAG	1325
Qy	1397	GAATGCAATGGGAAACAATTGCCATGAACGAGCTGAAGGTGGCCACGGCCCTGACCT	1456
Db	1326	GAACTGCATCGGCGACAGAGTTGCCATGCTGAGAGTAAAGGTGGCCATTGCTGATTCT	1385
Qy	1457	GCTCCGCTTTGAGCTGCTGCTGATCCACAGATCCCAATCCCATTTGACAGACTTGT	1516
Db	1386	GCTCCACTTCAAGTAGTCTCAGACCTCACAGGCTCTCTGCTTCTCCAGCCACACTGT	1445
Qy	1517	GTTGAATCCAAAAATGAAATCCACTTGCCTCAGAGAGCTCCCTAACCTT	1569
Db	1446	CCTCAGACCCAGCATGAAATCTAATTGCACTGAAGAAACTCTTGAAGTT	1498

LOCUS 792 bp mRNA linear EST 12-JUN-2001  
DEFINITION 602841036P1 NCI\_CGAP\_Kid14 Mus musculus CDNA clone IMAGE:4975160  
5', mRNA sequence.  
ACCESSION BG973046  
VERSION BG973046  
KEYWORDS GI:14360683  
EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 792)  
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
TITLE  
COMMENT unpublished  
Contact: Robert Strausberg, Ph.D.

```

FEATURES
source
      . location/Qualifiers
      1. .792
      /organism="Mus musculus"

```

```

/mol_type="mRNA"
/strain="PVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4975160"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1ib="NCI CGAP Kid14"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

```

Query Match	20.3%	Score 522.6;	DB 12;	Length 792;
Best Local Similarity	80.4%;	Pred. No. 1.1e-46;		
Matches 612;	Conservative 0;	Mismatches 149;	Indels 0;	Gaps 0;

Qy	531	CGAGTATGCTGGA	CAAA	TGGG	AAGACT	CTCTTG	GGC	CAGG	ATTCCC	CTCTG	GAGG	CTTT	590		
Db	32	CGATTGATGCTAGA	CAAA	TGGG	AAGCGG	CTGG	CTGAT	CAGG	ACTC	CTTAT	AGAT	CTTT	91		
Qy	591	CAGCAGTCTCCT	TGATGA	CCCT	TGAC	CACT	ATC	TGA	AGT	GTGC	TT	CAGCC	ATCAGGGC	650	
Db	92	CAACACATCTCCT	TATAG	ACCCT	AGAC	ACTGT	CATGA	AAAT	GTGC	CTT	CAGCC	CAAGGGC	151		
Qy	651	AGCATCCAGT	TGACAG	GAAT	TTCA	GTCT	CATAC	TAC	AGG	CCATT	AGT	AGC	CTGA	CAAC	710
Db	152	AGTGTTCAGG	TAGATGA	GAAT	TACAG	AGC	TCTAC	CCAGG	CCATT	GGG	CACTGA	ATTAAC	211		
Qy	711	CTGTTTTT	CCCG	TGTGA	GGAATG	CC	TTTCA	CCAG	AATG	ACAC	CACT	TACAG	CC	CTGACC	770
Db	212	CTATTTCA	CTCCCG	TGTGA	GGAAT	ATCT	TTTCA	TACAG	AATG	ATAC	CACT	TACAA	CTTT	CT	271
Qy	771	TCTGCTG	CCGCTG	GACAC	ACCG	CGCTG	CCAG	CTGG	CCCAT	CAG	CA	CAG	CA	CAAG	830
Db	272	TCCAATG	CCGCTTG	GGC	CAAA	CAAG	CTTGT	CACTT	GCC	ATG	ATC	AC	CAG	ATG	331
Qy	831	ATCCAACTG	AGAGG	CTCA	ACTAC	AGAA	GAGG	GGG	AGCTG	GAG	AAG	ATCA	AG	GGAG	890
Db	332	ATCAAGCTG	AGAA	GAGAT	CAG	CTG	CAG	ATG	AGG	GAG	AGCTG	GAA	AA	AGATCA	931
Qy	891	AGGCATTG	GATTTCT	GATAT	CTCT	CTTG	GCC	AAAT	GAG	AATG	GAG	ATCT	TG	950	
Db	392	AGACGTTG	ATTTCT	GATAT	CTTT	TAAT	TG	CCAG	AATG	AG	ATG	GG	CA	GATG	451
Qy	951	TCAGACA	GAGAC	CTCCG	TGCTG	AG	TGAC	AG	CTTCA	TGTT	GAG	GC	CA	CA	1010
Db	452	TCTGACA	GAGAC	CTA	CGT	CTG	AG	TGAC	ACA	TTC	ATG	TTG	AG	GG	511
Qy	1011	GCCAGT	GGAATCT	CTG	GATCT	CTAT	GTCT	TG	GG	CCAC	AC	CCCA	AG	CA	1070
Db	512	GCCAGT	GAGTCT	CTG	GAATCT	TAT	GTCT	TG	GG	CCAC	AC	CCCT	GAT	CA	571
Qy	1071	TGCCGG	AGAGAT	CCAC	AG	CTCT	CG	TGG	TGAT	GAG	CCCT	CA	CTG	AA	1130
</															

RESULT 11	
BI148819	
LOCUS	844 bp
BI148819	mRNA
	linear
	EST 05-JUL-2001









```
/clone="IMAGE:5064889"  
/lab_host="DH10B (T1 phage-resistant)"  
/clone_lib="NCI_CGAP_Kid14"  
/note="Organ: kidney; Vector: pCMV-SPORT6; Site_1: NotI;  
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.75 kb. Constructed by Life  
Technologies. Note: this is a NCI_CGAP library." "
```

BASE COUNT 202 a 215 c 184 g 205 t  
ORIGIN

```
Query Match 18.7%; Score 482.2; DB 12; Length 806;  
Best Local Similarity 77.5%; Pred. No. 2.1e-42;  
Matches 599; Conservative 0; Mismatches 168; Indels 6; Gaps 1;  
  
QY 13 GATCCAGCAGGTGCTGCACATGAGTGTCTGTGTGAGCCCCAGCAGACTCTGGGTG 72  
DB 1 GATCCAGAGAGCTGTGTATCATGAGTGCCTGTGCTGAGCTCCATCAGATTCCAGGAA 60  
  
QY 73 ATGCTCTGGAATCCTCCAAGCGGCTCCCTGCTCATTTCTTCTGCTGCTGATCAAG 132  
DB 61 GCATCTGTAGTACCTTCAAGTAGCCTGTGTGCTCAGCCTGCTCCTGCTCTTCAAGA 120  
  
QY 133 CAGTTCACTCTACCTGCAAGGCAAGTGGTGTCTCAAGCCCTCCAGCAGTTCCCGTCC 192  
DB 121 CAGCCCACTCTACCTGCAAGGCAATGCTACTCAGCAGTACTCAGCAGTTCCCATCCC 180  
  
QY 193 CTCCCTCCCACTGCTGCTTGGGGCATCCAGAGCTCCAAGAGCAGACAGGAGCTACAAC 252  
DB 181 CACCTTCTCACTGGCTCTTGGACACA-----AGATCTTAAGAGACAGGAGCTTCAAG 234  
  
QY 253 GGATTCAGAATGGGTGAGACATTCCCAAGTGCCTGCTCATTTGGCTATGGGAGGCA 312  
DB 235 ATATTCTAAGTAAAGATTTCCTCAAGTGCCTGTCCACAGTGGCTCTGGGGAAGCA 294  
  
QY 313 AAGTTCGTGTCAGCTCTATGACCTGACTATATGAAGTGAATTCTGGGAGATCAGACC 372  
DB 295 AAGTGCAGATTCAAGTATGACCTGACTACATGAAGTGAATTCTGGGAGATCAGACC 354  
  
QY 373 CGAATCCCATGTTCTCTACAGATTCTGCTCCATGATGGGTACGGCTTCTCTGT 432  
DB 355 CAAAGCTAATGTTCTCTACAGATTCTAGCTCCCTGATGGGCGTGTCTTCTTGC 414  
  
QY 433 TGAATGGGAGACATGTTCCAGCATGACGAGTGTGACCCAGCCTTCCACTATGACA 492  
DB 415 TGGATGGACAGACATGTTTCAAGACCGAGATGTGACCCAGCCTTCCACTATGACA 474  
  
QY 493 TCTGAAAGCCCTATGTGGGGCTCATGSCAGACTCTGTACAGATGATGCTGGACAAATGGG 552  
DB 475 TTCTGAAGCCTTATACGGAATCATGSCAGACTCTGTGTATATGCTGATTAATGGG 534  
  
QY 553 AAGAGCTCCTTGGCCAGGATTCCTCTGAGGTCTTTCAGCAGCTCTCTTGATGACCC 612  
DB 535 AACAGATTGTGGCCAGGATTCACCTGAGATCTTTCAGACATCACCTTGATGACCT 594  
  
QY 613 TGGACACCATCATGAAGTGTGCTTCAAGCATCAGGAGCAGCATCCAGGTGACAGGAATT 672  
DB 595 TGGACACCATCATGAAGTGTGCTTCAAGCATCAGGAGCAGGTTCAGTTGACAGAAAT 654  
  
QY 673 CTGAGTCTTACATACAGGCAATGAGCTGAAACAACCTGTTTTTCCCGTGAAGA 732  
DB 655 ACAAGTCTTATATCCAGGCAATGAGGAGCTGAAAGATCTGTTTTTCCCGTGAAGA 714  
  
QY 733 ATGCTTTTCAACGAAATGACACCATCTACAGCCTGACCTTGTGCGCGCTGG 785  
DB 715 CATTTCTTCAACGAAATGACATCTTACAGAGTGTCTTATATGCTGACAGG 767
```

Search completed: February 15, 2004, 05:29:04  
Job time : 5307 secs